

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2019 年 11 月 28 日 (28.11.2019)



(10) 国际公布号
WO 2019/223552 A1

(51) 国际专利分类号:
G06F 17/00 (2019.01)

(21) 国际申请号: PCT/CN2019/086374

(22) 国际申请日: 2019 年 5 月 10 日 (10.05.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

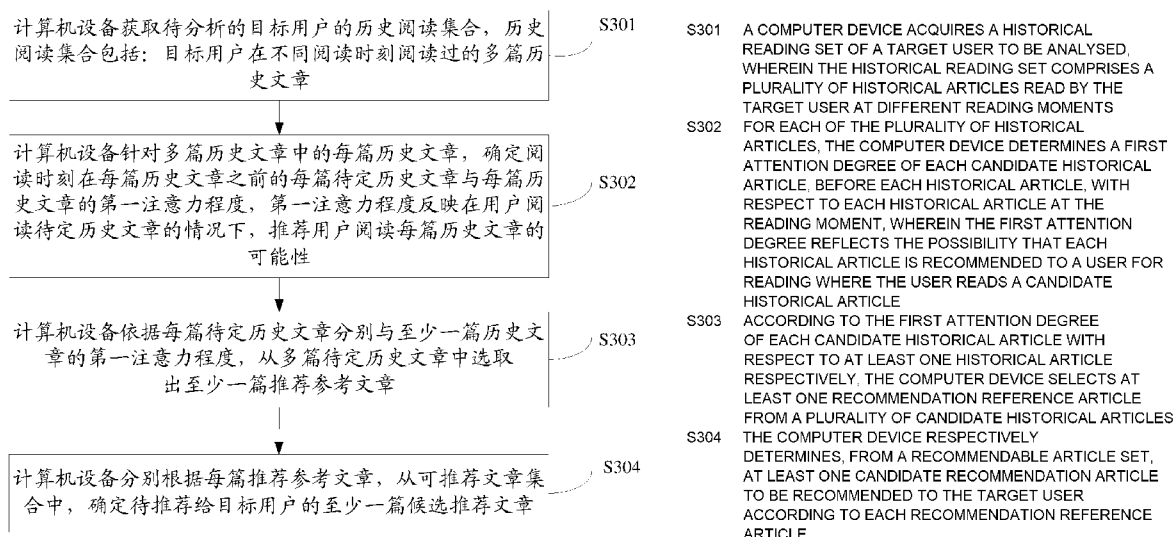
(30) 优先权:
201810515869.8 2018年5月25日 (25.05.2018) CN

(71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。

(72) 发明人: 刘毅(LIU, Yi); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 胡澜涛(HU, Lantao); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 张博(ZHANG, Bo); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 夏锋(XIA, Feng); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 林乐宇(LIN, Leyu); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 冯喆(FENG, Zhe); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 陈磊(CHEN, Lei); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 饶君(RAO, Jun); 中

(54) **Title:** ARTICLE RECOMMENDATION METHOD AND APPARATUS, AND COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 文章推荐方法、装置、计算机设备及存储介质



(57) **Abstract:** Provided are an article recommendation method and apparatus, and a computer device and a storage medium. The method comprises: a computer device acquiring a historical reading set of a target user to be analysed, wherein the historical reading set comprises a plurality of historical articles read by the target user at different reading moments; for each of the plurality of historical articles, the computer device determining a first attention degree of each candidate historical article, before each historical article, with respect to each historical article at the reading moment; according to the first attention degree of each candidate historical article with respect to at least one historical article respectively, the computer device selecting at least one recommendation reference article from

国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 刘书凯(LIU, Shukai); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。

丘志杰(QIU, Zhijie); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 孙振龙(SUN, Zhenlong); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 王良栋(WANG, Liangdong); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。

(74) 代理人: 北京康信知识产权代理有限公司(KANGXIN PARTNERS, P.C.); 中国北京市海淀区知春路甲48号盈都大厦A座16层, Beijing 100098 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

a plurality of candidate historical articles; and the computer device respectively determining, from a recommendable article set, at least one candidate recommendation article to be recommended to the target user according to each recommendation reference article. By means of the solution of the present application, the diversity of articles recommended to a user can be improved.

(57) 摘要: 本申请提供了一种文章推荐方法、装置、计算机设备及存储介质, 该方法包括: 计算机设备获取待分析的目标用户的历史阅读集合, 历史阅读集合包括: 目标用户在不同阅读时刻阅读过的多篇历史文章; 计算机设备针对多篇历史文章中的每篇历史文章, 确定阅读时刻在每篇历史文章之前的每篇待定历史文章与每篇历史文章的第一注意力程度; 计算机设备依据每篇待定历史文章分别与至少一篇历史文章的第一注意力程度, 从多篇待定历史文章中选取至少一篇推荐参考文章; 计算机设备分别根据每篇推荐参考文章, 从可推荐文章集中, 确定待推荐给目标用户的至少一篇候选推荐文章。本申请的方案可以提高向用户推荐文章的多样性。

文章推荐方法、装置、计算机设备及存储介质

本申请要求于 2018 年 5 月 25 日提交中国专利局、优先权号为 2018105158698、申请名称为“文章推荐方法、装置、计算机设备及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及数据处理技术领域，尤其涉及一种文章推荐方法、装置、计算机设备及存储介质。

背景技术

随着网络的不断发展，用户通过互联网中各种文章发布平台（如，即时通讯平台中的各种公众号等）所能阅读到的文章数量也日益增多。

为了吸引用户阅读文章发布平台的文章，文章发布平台会根据用户阅读过的历史文章，向用户推荐文章。但是目前为用户推荐文章时，都是根据用户阅读历史，向用户推荐内容上相似的文章。如，基于内容的推荐技术是根据用户历史阅读文章，确定与用户阅读过的文章的内容相似的文章，并将这些相似文章推荐给用户。

然而，由于为用户推荐的文章与用户历史阅读过的文章的内容大多相似，会导致文章推荐的多样性较差。

发明内容

有鉴于此，本申请提供了一种文章推荐方法、装置、计算机设备及存储介质，以提高文章推荐的多样性。

为实现上述目的，一方面，本申请提供了一种文章推荐方法，包括：

计算机设备获取待分析的目标用户的历史阅读集合，上述历史阅读集合包括：上述目标用户在不同阅读时刻阅读过的多篇历史文章；

上述计算机设备针对上述多篇历史文章中的每篇历史文章，确定阅读时刻在上述每篇历史文章之前的每篇待定历史文章与上述每篇历史文章的第一注意力程度，上述第一注意力程度反映在用户阅读上述待定历史文章的情况下，推荐用户阅读上述每篇历史文章的可能性；

上述计算机设备依据每篇上述待定历史文章分别与至少一篇历史文章的第一注意力程度，从多篇上述待定历史文章中选取出至少一篇推荐参考文章；

上述计算机设备分别根据每篇上述推荐参考文章，从可推荐文章集合中，确定待推荐给上述目标用户的至少一篇候选推荐文章。

又一方面，本申请还提供了一种文章推荐装置，应用在计算机设备中，包括：

历史获取单元，设置为获取待分析的目标用户的历史阅读集合，上述历史阅读集合包括：上述目标用户在不同阅读时刻阅读过的多篇历史文章；

第一注意分析单元，设置为针对上述多篇历史文章中的每篇历史文章，确定阅读时刻在上述每篇历史文章之前的每篇待定历史文章与上述每篇历史文章的第一注意力程度，上述第一注意力程度反映在用户阅读上述待定历史文章的情况下，推荐用户阅读上述每篇历史文章的可能性；

推荐参考确定单元，设置为依据每篇上述待定历史文章分别与至少一篇历史文章的第一注意力程度，从多篇上述待定历史文章中选取出至少一篇推荐参考文章；

候选推荐确定单元，设置为分别根据每篇上述推荐参考文章，从可推荐文章集合中，确定待推荐给上述目标用户的至少一篇候选推荐文章。

又一方面，本申请还提供了一种计算机设备，包括：

处理器和存储器；

其中，上述处理器用于执行上述存储器中存储的计算机程序；

上述存储器用于存储上述计算机程序，上述计算机程序至少用于：

获取待分析的目标用户的历史阅读集合，上述历史阅读集合包括：上述目标用户在不同阅读时刻阅读过的多篇历史文章；

针对上述多篇历史文章中的每篇历史文章，确定阅读时刻在上述每篇历史文章之前的每篇待定历史文章与上述每篇历史文章的第一注意力程度，上述第一注意力程度反映在用户阅读上述待定历史文章的情况下，推荐用户阅读上述每篇历史文章的可能性；

依据每篇上述待定历史文章分别与至少一篇历史文章的第一注意力程度，从多篇上述待定历史文章中选取出至少一篇推荐参考文章；

分别根据每篇上述推荐参考文章，从可推荐文章集合中，确定待推荐给上述目标用户的至少一篇候选推荐文章。

又一方面，本申请还提供了一种存储介质，上述存储介质中存储有计算机程序，上述计算机程序被运行时，执行上述上述上述文章推荐方法

经以上可知，在本申请实施例中，对于用户阅读过的历史文章，会分析用户在阅读该历史文章之前阅读的其他历史文章与该历史文章的注意力程度，由于两篇文章之间的注意力程度可以反映出用户阅读一篇文章的情况下，推荐用户阅读另一篇文章的可能性，因此，依据确定出的各篇文章与其他文章之间的注意力程度，可以从阅读历史集合中确定出表征向目标用户推荐文章的可能性较低的历史文章，而该历史文章的类型及内容便不属于用户经常阅读到的文章类型及文章内容，因此，以该历史文章作为推荐参考文章，向用户推荐候选推荐文章，有利于提高确定出的候选推荐文章的多样性，进而有利于提高向用户推荐文章的多样性。

附图说明

为了更清楚地说明本申请实施例或相关技术中的技术方案，下面将对实施例或相关技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的实施例，对于本领域普通技术人员来

讲，在不付出创造性劳动的前提下，还可以根据提供的附图获得其他的附图。

图1示出了本申请实施例的文章推荐方法所适用的一种系统组成架构示意图；

图2示出了本申请实施例的文章推荐方法所适用的一种计算机设备的组成架构示意图；

图3示出了本申请实施例的文章推荐方法的一种流程示意图；

图4示出了本申请实施例的文章推荐方法中训练注意力模型的一种流程示意图；

图5示出了本申请实施例中基于LSTM输出文章的文章向量的示意图；

图6示出了本申请实施例中确定推荐参考文章的一种实现原理示意图；

图7示出了本申请实施例的文章推荐方法又一个实施例的流程示意图；

图8示出了本申请实施例的文章推荐方法的一种实现原理示意图；

图9示出了本申请实施例的另一种文章推荐方法的流程示意图；

图10示出了本实施例的一种可选的界面显示图；

图11示出了本实施例的另一种可选的界面显示图；

图12示出了本实施例的又一种可选的界面显示图；

图13示出了本申请实施例的文章推荐装置的一种组成架构示意图；

图14示出了本申请实施例的文章推荐装置的又一种组成架构示意图。

具体实施方式

本申请实施例的文章推荐方法适用于各种文章发布平台，以确定向文章发布平台中不同用户所需要推荐的文章，提高向用户推荐文章的多样性。如，新闻媒体平台针对不同用户推荐新闻文章；社交平台或者社交平台接入的公众号向用户推荐文章等。

为了便于理解，参见图1，其示出了本申请的一种文章推荐系统的组成架构示意图。

由图1可知，该文章推荐系统可以包括：计算机设备101以及至少一台文章发布服务器102。

其中，文章发布服务器102，用于发布可供用户阅读的多篇文章，并记录不同用户在不同时刻阅读过的文章。为了便于区分，将用户阅读过的文章称为历史文章。

该计算机设备101，用于分别基于每个用户阅读过的多篇历史文章，确定需要向该用户推荐的推荐文章。

该计算机设备101还可以将确定出的待推荐给用户的推荐文章的信息发送给文章发布服务器。

可选的，该文章推荐系统还可以包括数据存储服务器103，该数据存储服务器可以存储至少一台文章发布服务器所产生的用户数据，如用户画像，以及用户的阅读历史记录等等。其中，用户画像可以包括用户的性别、年龄、学历以及所处地区等用户属性信息。阅读历史记录可以包括：用户阅读过的文章以及文章相关的来源等等信息。

相应的，计算机设备101可以从数据存储服务器103获取用户的阅读历史记录，以根据用户阅读历史记录，确定需要向该用户推荐的推荐文章。

需要说明的是，计算机设备101所依据的用户阅读过的历史文章可以是来源于某一个文章发布平台中的一台或多台文章发布服务器，也可以是来源于多个文章发布平台的文章发布服务器。

其中，对于某个文章发布平台而言，计算机设备101可以是独立于文章发布服务器单独设置的服务器或者终端；也可以是与文章发布服务器为同一台服务器，也就是说，该计算机设备具备文章发布以及确定向用户推荐文章的功能。

为了便于理解该计算机设备，如，参见图2，其示出了本申请的计算机设备的一种组成示意图。

在图2中，该计算机设备200可以包括：处理器201、存储器202、通信接口203、输入单元204和显示器205和通信总线206。

处理器 201、存储器 202、通信接口 203、输入单元 204、显示器 205、均通过通信总线 206 完成相互间的通信。

在本申请实施例中，该处理器 201，可以为中央处理器（Central Processing Unit，CPU），特定应用集成电路（application-specific integrated circuit，ASIC），数字信号处理器（DSP）、专用集成电路（ASIC）、现成可编程门阵列（FPGA）或者其他可编程逻辑器件等。

该处理器可以调用存储器 202 中存储的程序，处理器可以执行以下图 3 至图 9 中服务器侧所执行的操作。

存储器 202 中用于存放一个或者一个以上程序，程序可以包括程序代码，程序代码包括计算机操作指令，在本申请实施例中，该存储器中至少存储有用于实现以下功能的程序：

获取待分析的目标用户的历史阅读集合，历史阅读集合包括：目标用户在不同阅读时刻阅读过的多篇历史文章；

针对每篇历史文章，依据历史文章的内容特征以及历史阅读集合中阅读时刻在历史文章的阅读时刻之前的至少一篇待定历史文章的内容特征，分别确定每篇待定历史文章与历史文章的第一注意力程度，第一注意力程度反映在用户阅读待定历史文章的情况下，推荐用户阅读历史文章的可能性；

依据得到的多篇待定历史文章分别与至少一篇历史文章的第一注意力程度，从多篇待定历史文章中，选取出触发向目标用户推荐文章的可能性相对较低的至少一篇待定历史文章作为推荐参考文章；

分别根据每篇推荐参考文章的内容特征，从可推荐文章集合中，确定待推荐给目标用户的至少一篇候选推荐文章。

其中，存储器 202 可包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、以上所提到的程序，以及至少一个功能（比如声音播放功能、图像播放功能等）所需的应用程序等；存储数据区可存储根据计

计算机设备的使用过程中所创建的数据以及接收到的待处理数据，比如，用户阅读记录等。

此外，存储器 202 可以包括高速随机存取存储器，还可以包括非易失性存储器，例如至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

该通信接口 203 可以为通信模块的接口。

本申请中该服务器还可以包括输入单元 205，如，键盘等等。

该显示器 204 包括显示面板。在一种可能的情况中，可以采用液晶显示器 (Liquid Crystal Display, LCD)、有机发光二极管 (Organic Light-Emitting Diode, OLED) 等形式来配置显示面板。

当然，图 2 所示的计算机设备结构并不构成对本申请实施例中计算机设备的限定，在实际应用中计算机设备可以包括比图 2 所示的更多或更少的部件，或者组合某些部件。

结合以上共性，对本申请实施例的一种文章推荐方法进行介绍。如，参见图 3，其示出了本申请实施例一种文章推荐方法的一种流程示意图，本实施例从计算机设备侧进行介绍。该过程包括：

S301，获取待分析的目标用户的历史阅读集合。

其中，历史阅读集合包括：该目标用户在不同阅读时刻阅读过的多篇历史文章。

如，历史阅读集合包括按照阅读先后顺序排序的多篇历史文章，这样，根据该多篇历史文章的先后顺序，便可以确定该多篇历史文章对应的阅读时刻的先后顺序。

又如，该历史阅读集合可以包括多篇目标用户阅读过的多篇历史文章以及每篇历史文章对应的阅读时刻。

可以理解的是，针对不同的用户，需要结合不同用户的历史阅读集合，来确定适合推荐给该用户的文章，在本申请实施例中，为了便于描述，将当前需要确定待推荐文章的用户称为待分析的目标用户。同时，为了与推荐给用户的文章进行区分，将用户阅读过的文章称为历史文章。

S302, 针对每篇历史文章, 依据该历史文章的内容特征, 计算机设备从历史阅读集合中阅读时刻在该历史文章的阅读时刻之前的至少一篇待定历史文章。

可以理解的是, 对于历史阅读集合中的任意一篇历史文章, 可以认为是在该目标用户通过阅读该历史文章之前, 由于阅读其他历史文章而阅读到该历史文章。而此处的其他历史文章与该历史文章之间的关系可以表征出其他历史文章与该历史文章之间的聚焦程度, 即, 以其他历史文章为基准, 推荐用户阅读该历史文章的可能性, 因此, 本申请需要确定用户在阅读历史文章之前阅读的其他历史文章, 并分析相应的历史文章之间的关系。

如, 历史文章集合中包括文章 A、文章 B 和文章 C, 且这三篇文章的顺序表征了这三篇文章被用户阅读的阅读先后顺序, 由此可知, 文章 A 和文章 B 是在阅读文章 C 之前阅读的, 那么文章 C 有可能是由于用户阅读了文章 A 和文章 B 中的一篇或者两篇文章后, 选择阅读的该文章 C, 因此, 需要确定文章 C 之前阅读的文章 A 与文章 B, 并分析文章 A、文章 B 分别与文章 C 的关系, 以得到由文章 A 触发用户阅读文章 C 的可能性, 以及由文章 B 触发用户阅读文章 C 的可能性。

其中, 为了便于区分, 将该历史文章之前阅读过的, 用于分析与该历史文章之间关系的其他历史文章均称为待定历史文章。

在本申请实施例, 可以将该历史文章之前阅读过的每篇待定历史文章均作为待定历史文章, 并执行后续分析待定历史文章与该历史文章之间关系的操作。可选的, 考虑到与在该历史文章之前阅读过的文章数量较大, 为了降低数据处理量, 又不影响后续分析结果, 可以将该历史文章之前阅读到的设定数量篇历史文章作为待定历史文章。例如, 从该历史文章之前阅读的其他历史文章中, 随机选取设定数量篇历史文章, 以将选取出的设定篇历史文章作为待定历史文章, 并执行后续操作。

考虑到与该历史文章的阅读时刻较为接近的其他历史文章, 对于用户能否阅读到该历史文章的影响性较大, 因此, 可以从阅读时刻处于该历史文章的阅读时刻之前, 且阅读时刻与该历史文章的阅读时刻之间的时长不

超过预设时长的多篇历史文章中，随机选取预设数量篇历史文章作为待定历史文章或者选取与该历史文章的阅读时刻最近的多篇历史文章作为待定历史文章。

需要说明的是，考虑到历史阅读集合中，会存在某些历史文章之前阅读的文章的数量不足预设数量，例如，历史阅读集合中阅读时刻处于第二位的历史文章之前仅仅有一篇文章，因此，在该种情况中，历史文章对应的待定历史文章的数量可以少于该预设数量。

S303，针对每篇历史文章，依据该历史文章的内容特征以及该历史文章对应的至少一篇待定历史文章的内容特征，计算机设备分别确定每篇待定历史文章与该历史文章的第一注意力程度。

其中，该注意力程度也可以称为聚焦程度，反映的是阅读一篇文章的情况下，推荐用户阅读另一篇文章的可能性。同时，注意力程度也可以反映出在用户一篇文章的情况下，通过点击等方式选择阅读另一篇文章的可能性。正是由于注意力程度的以上含义，其也被称为从一篇文章到另一篇文章的注意力程度或者聚焦程度。

在本申请实施例中，注意力程度可以是一个分值、概率、等级等等，在此不加以限制。

其中，为了便于与后续其他文章之间的注意力程度进行区分，将待定历史文章与历史文章之间的注意力程度称为第一注意力程度。相应的，该第一注意力程度用于反映在用户阅读该待定历史文章的情况下，推荐用户阅读该历史文章的可能性。在该步骤 S303 中，计算机设备需要分别确定每篇待定历史文章各自与该历史文章的第一注意力程度。

可以理解的是，对于任意文章，文章的内容特征由该文章的标题及文章的正文中的字词决定，文章的内容特征可以反映出文章类型、文章的主题思想及文章所涉及到的具体内容等等。因此，对于任意两篇文章，可以依据这两篇文章各自的内容特征，来分析从其中一篇文章到另一篇文章的注意力程度。

在一种可能的实现方式中，对于任意两篇文章，基于这两篇文章各自

的内容特征，计算机设备可以分析这两篇文章的内容特征的相似性，并依据内容特征的相似性，确定这两篇文章之间的注意力程度。如，将这两篇文章的内容特征的相似性，作为这两篇文章之间的注意力程度。

相应的，计算机设备依据待定历史文章与历史文章的内容特征，确定待定历史文章与历史文章的相似性，确定待定历史文章与历史文章之间的注意力程度。

特别的，大部分情况需要确定一个文章集合中各篇文章 M 与某一篇文章 N 的注意力程度的情况，而且可能需要比较不同文章集合中的文章 M 与该篇文章 M 的注意力程度之间的大小关系。在该种情况下，为了便于比较各篇文章 M 与该篇文章 N 之间的注意力程度，对于每个文章集合，可以在分别确定出该文章集合中每篇文章 M 与该篇文章 N 的内容特征的相似性之后，根据该文章集合中各篇文章 M 各自对应的相似性，分别对该文章集合中每篇文章 M 与文章 N 的相似性进行归一化，并将每篇文章 M 与该文章 N 之间的相似性对应的归一化结果作为反映注意力程度的评价指标。

其中，计算机设备按照上面的方式，对一篇文章 M 与文章 N 的内容特征的相似性进行归一化之后，归一化结果都是一个大于等于零小于 1 的数值，通过该数值可以反映从该文章 M 触发点击文章 N 的可能性。

在又一种可能的实现方式中，计算机设备可以预先训练得到用于确定注意力程度的注意力模型。该注意力模型为利用多个用户的历史文章样本集合以及每个历史文章样本集合中每篇历史文章样本各自对应的已标注的注意力分数序列训练得到的。其中，为了便于区分，将训练注意力模型的历史文章称为历史文章样本。

其中，对于一个用户而言，该用户的历史文章样本对应的注意力分数序列包括：用户阅读该历史文章样本之前阅读的多篇历史文章样本分别与该历史文章样本的注意力分数。如，通过从文章发布平台的服务器中获取用户的历史阅读日志，得到用户阅读不同文章时，通过点击等方式选择阅读的其他文章，从而统计出从各篇文章触发阅读其他文章的概率，并将相

应的概率作为注意力分数。

相应的，在该步骤 S303 中，计算机设备可以依据历史文章的内容特征以及该历史阅读集合中阅读时刻在该历史文章的阅读时刻之前的至少一篇待定历史文章的内容特征，并利用预先训练得到的注意力模型，分别确定每篇该待定历史文章与该历史文章的注意力分数。

可选的，文章的内容特征可以通过文章的文章向量表示。相应的，为了确定某一篇或者多篇文章 M 到另一篇文章 N 的注意力分数，可以先分别确定每篇文章（如，文章 M 及文章 N）的内容特征，确定出该文章的文章向量。然后，将各篇文章 M 以及文章 N 的文章向量输入到该注意力模型中，以分别输出每篇文章 M 到文章 N 的注意力分数。

可以理解的是，为了便于后续比较不同待定历史文章推荐出其他历史文章的可能性的的大小，该注意力模型输出的各个待定历史文章与历史文章之间的注意力程度（如注意力分数）同样可以为归一化后的结果。

其中，依据历史样本集合训练注意力模型的过程可以有多种，本申请对此不加以限制。为了便于理解，以注意力模型为通过对深度神经网络进行训练得到的为例进行说明。如，参见图 4，其示出了训练注意力模型的一种流程示意图。

在 S401 部分中，计算机设备获取多个用户的历史文章样本集合以及每个历史文章样本集合中各篇历史文章样本各自对应的已标注的注意力分数序列。

如，对于某个用户的历史文章样本集合中包括用户先后阅读的文章 S1，文章 S2、文章 S3 和文章 S4，其中，对于文章 S4 对应的已标注的注意力分数序列可以包括：该文章 S4 与文章 S3 的注意力分数，文章 S4 与文章 S2 的注意力分数、该文章 S4 与文章 S1 的注意力分数。而对于文章 S3 而言，其对应的已标注的注意力分数序列可以包括：文章 S3 与文章 S2 的注意力分数，文章 S3 与文章 S1 的注意力分数。对于文章 S2 和 S1 这种阅读时刻处于其阅读时刻之前的文章仅有一篇或者没有文章的情况，可以不设置注意力分数序列。

可以理解的是, 历史文章样本对应的已标注的注意力分数序列中每个注意力分数为依据该注意力分数序列中各个注意力分数归一化之后的结果。

在步骤 S402 部分中, 计算机设备分别将各篇历史文章样本输入到深度神经网络中进行循环训练, 并依据各篇历史文章样本各自的注意力分数序列对深度神经网络模型进行循环训练, 直至深度神经网络输出的各篇历史文章对应的实际注意力分数序列分别与相应的已标注的注意力分数序列之间的准确度满足预设要求。

如, 以历史样本集合 G 中, 用户在 N 时刻阅读的历史文章样本 S_m 对应的已标注的注意力分数序列为例进行说明, 历史文章样本 S_m 对应的已标注的注意力分数序列为该历史文章样本 S_m 分别与历史文章样本序列 S 中多篇历史文章样本之间的注意力分数, 其中, 历史文章样本序列 S 为该历史样本集合 G 中在 N 时刻之前被阅读的多篇历史文章样本所组成的序列。

相应的, 历史文章样本 S_m 对应的已标注的注意力分数序列可以表示为: $A_m^* = \{a_1^*, a_2^*, \dots, a_{m-1}^*\}$, 其中, a_j^* 表示历史文章样本 S_m 与历史文章样本序列 S 中历史文章样本 S_j 之间的注意力分数, j 表示从 1 到 $m-1$ 的自然数, $m-1$ 表示历史文章样本序列 S 中历史样本的总个数。

计算机设备将该历史文章样本 S_m 及其对应的已标注的注意力分数序列 A_m^* 输入到该深度神经网络模型中, 可以该深度神经网络模型可以输出该历史文章样本 S_m 对应的实际注意力分数序列 A_m , 该实际注意力分数序列 A_m 包括该深度神经网络模型输出的该历史文章样本 S_m 分别与历史文章样本序列 S 中各篇历史文章样本的注意力分数。

这样, 通过比较已标注的注意力分数序列 A_m^* 与实际注意力分数序列 A_m 的差距, 可以分析该深度神经网络模型的准确度是否符合要求。可以依据预设的损失函数, 以及多个历史文章样本分别对应的已标注的注意力分数序列和实际注意力分数序列, 并采用梯度下降法来进行优化, 直至迭代收敛。如, 可以定义如下的损失函数 (即公式一):

$$L(A, A^*) = \sqrt{\sum_{m=1}^n (A_m - A_m^*)^2} \quad (\text{公式一});$$

其中, m 取值不同的是, A_m^* , A_m 分别表示不同历史文章样本对应的已标注的注意力分数, 与实际注意力分数, n 为训练用的历史文章样本的总个数。

相应的, 假设优化目标为 $\min L(A, A^*)$, 则计算机设备可以使用梯度下降的方法对该优化目标进行优化, 直至最终迭代收敛, 则训练结束。训练出的深度神经网络模型可以作为注意力模型。

其中, 训练得到的注意力模型同样是根据不同文章的内容特征, 来计算文章之间的相似度, 并基于相似度, 确定文章的注意力程度。

举例说明, 以需要确定一个文章序列 (集合) $D = \{d_1, d_2, \dots, d_i, \dots, d_{t-2}, d_{t-1}\}$ 中任一篇文章 d_i 分别与文章 d_t 的注意力程度为例进行说明。其中, d_i 表示文章序列 D 中任意一篇文章, i 为从 1 到 $t-1$ 的自然数, $t-1$ 为文章序列 D 中文章的总数量。为了确定文章序列 D 中任一篇文章 d_i 与文章 d_t 之间的注意力程度, 注意力模型中该文章序列 D 中文章 d_i 与文章 d_t 的注意力分数的函数关系可以表示为如下公式二:

$$a_i = \frac{\exp(e_i)}{\sum_{j=1}^{t-1} \exp(e_j)} \quad (\text{公式二});$$

其中,

$$e_i = F(Wh_{d_i} + Uh_{d_t} + b) \quad (\text{公式三});$$

其中, h_{d_i} 为文章 d_i 的文章向量, h_{d_t} 为文章 d_t 的文章向量。 W , U , b 均为注意力模型中设定的参数, 这些参数的参数值在训练过程中确定。 F 表示一种设定的函数关系, 该函数关系同样可以在训练过程中确定。

$$\text{如, } F(x) = \tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}。$$

其中, e_i 求取的是文章 d_i 与文章 d_t 之间文章向量的相似性, 而通过公式二进行归一化之后, 可以得到文章 d_i 与文章 d_t 的注意力分数。

其中，对于任意一篇文章，基于文章的内容特征确定该文章的文章向量的方式也可以有多种。可选的，计算机设备可以是将该文章中的内容进行分词，并将该文章分词出的多个词语输入到预先训练得到的文章向量模型中，以通过文章向量模型输出该文章的文章向量。

如，以文章向量模型为长短期记忆网络（Long Short-Term Memory, LSTM）为例进行说明。为了确定文章 d 的文章向量，首先，将文章的内容进行分词，假设文章 d 分词出 N 个词，则文章可以表示为 $d = \{w_1, w_2, \dots, w_N\}$ ，其中， w_i 表示文章 d 中分词出的第 i 个分词， i 为从 1 到 N 的自然数。然后，将这 N 个分词依次输入到该 LSTM 网络模型中，从而可以通过该 LSTM 网络模型最终输出该文章 d 的文章向量。

为了便于理解，可以参见图 5，其示出了 LSTM 网络模型将文章的文章向量的一种示意图。

由该图 5 可以看出，LSTM 网络模型中的输入为文章分词出的各个分词，最终输出为文章的文章向量为 h_i 。

S304，依据得到的多篇该待定历史文章分别与至少一篇历史文章的注意力程度，从该多篇待定历史文章中，选取出触发向该目标用户推荐文章的可能性相对较低的至少一篇待定历史文章作为推荐参考文章。

可以理解的是，在目标用户的历史阅读集合中，如果目标用户阅读某篇历史文章的情况下，触发向目标用户推荐其他文章或者说触发目标用户点击阅读其他文章的可能性较低，则说明在当前时刻之前，基于该篇历史文章向目标用户推荐文章的概率较低。而由于该篇历史文章存在于历史阅读集合中，所以该篇历史文章的内容或者其他相关信息是用户感兴趣的，但是却很少获取到相关信息的文章。在该种情况下，如果后续基于该篇历史文章向该目标用户推荐文章，必然有利于扩展用户所阅读到的文章，增加文章推荐的多样性。

举例说明，假设目标用户的历史阅读集合中，存在属于时尚类、文艺类、新闻资讯类、科学类等类型的文章，而对于属于同一类型或者内容相似的两篇或者多篇文章而言，通过前面步骤，确定出属于同一类型的两篇

或多篇文章之间的注意力程度会相对较高。因此，如果确定某篇待定历史文章与多个历史文章之间的注意力程度相对较高，则说明该篇待定历史文章触发推荐文章的可能性高，而且历史阅读集合中与该待定历史文章的类型或者内容相似的文章也相对较多；反之，如果某篇待定历史文章与其他文章之间的注意力程度较低，则说明目标用户阅读过该篇待定历史文章相关类型或者内容的文章较少，因此，可以以该篇待定历史文章为推荐参考文章，以便增加向该目标用户推荐文章的多样性。

可以理解的是，由于每篇历史文章均对应的一个或者多个待定历史文章，因此，在该步骤 S303 中可以得到多篇待定历史文章。同时，一篇历史文章可能会作为一篇或者多篇其他历史文章的待定历史文章，相应的，一篇待定历史文章可能会分别确定出与多篇历史文章的注意力程度，这样，针对每篇待定历史文章，可以得到该待定历史文章与至少一篇历史文章的注意力程度。因此，针对每篇待定历史文章，综合该待定历史文章与对应的各篇历史文章之间的注意力程度，可以得到基于待定历史文章向目标用户推荐文章的可能性。

如，对于文章 4，可能会分别将文章 3、文章 2 和文章 1 作为待定历史文章，并分别确定文章 3、文章 2 以及文章 1 与文章 4 之间的注意力程度，同时，对于文章 3 也可能会将文章 2 和文章 1 作为待定历史文章，并分别确定出文章 2 和文章 1 与文章 3 的注意力程度。相应的。对于文章 3，仅仅确定出文章 3 与文章 4 之间的注意力程度，而该注意力程度可以用于反映基于文章 3 触发向用户推荐文章的可能性。而对于文章 2，则依据文章 2 与文章 4 之间的注意力程度，以及文章 2 与文章 3 的注意力程度，来综合确定基于该文章 2，触发向用户推荐文章的可能性。

其中，对于任意一篇待定历史文章，确定基于该待定历史文章向用户推荐文章的可能性的方式可以有多种。

如，在一种实现方式中，针对每篇待定历史文章，根据该待定历史文章分别与至少一篇历史文章（该待定历史文章对应的至少一篇历史文章）的注意力程度，确定该待定历史文章与该至少一篇历史文章的平均注意力

程度。相应的，该平均注意力程度可以反映基于该待定历史文章，向用户推荐文章的可能性，其中，该平均注意力程度越低，则触发向用户推荐文章的可能性也越低。

计算机设备可以从多篇该历史文章对应的多篇待定历史文章中，选取出平均注意力程度较低的至少一篇待定历史文章作为推荐参考文章，以便后续基于该推荐参考文章，确定需要向该目标用户推荐的文章。如，选取平均注意力程度最低的待定历史文章作为推荐参考文章。

例如，在对多篇待定历史文章中的每一篇待定历史文章都确定出一个第一注意力程度之后，可以对每一篇待定历史文章的第一注意力程度进行排序并获得排序结果。若是对第一注意力程度由大到小进行排序，则排序结果中越靠前的第一注意力程度对应的待定历史文章所对应的类型或内容是用户经常阅读的，而排序结果中越靠后的第一注意力程度对应的待定历史文章的类型或内容是用户不常阅读到的。此时可以根据需要确定推荐参考文章，如获取排序结果中最后一个第一注意力程度对应的待定历史文章，并将该待定历史文章作为推荐参考文章。则可以根据推荐参考文章向用户推荐不常阅读的文章类型。具体选择的推荐参考文章的数量可以为用户所阅读的多篇历史文章的一定百分比，如 30%。

为了便于理解，以基于待定历史文章的平均注意力程度，计算机设备从多篇待定历史文章中选取推荐参考文章为例，对本申请从历史文章集合中确定出推荐参考文章的过程进行举例说明。如，参见图 6，其示出了确定推荐参考文章的一种实现原理示意图。

在图 6 中以目标用户的历史阅读集合中包括文章 d1、文章 d2、文章 d3、文章 d4、文章 d5 和文章 d6 为例。在图 6 中以每篇文章选取在该历史文章之前阅读的三篇文章作为待定历史文章为例，同时，为了便于描述，选取与文章的阅读时刻相对较近的三篇文章作为待定历史文章。

如，图 6 中“待定历史文章选取”部分中，文章 d6 对应的待定历史文章包括：文章 d5、文章 d4 和文章 d3。文章 d5 对应的待定历史文章包括：文章 d4、文章 d3 和文章 d2。文章 d4 对应的待定历史文章包括：文

章 d3、文章 d2 和文章 d1。由于在文章 d3 和文章 d2 之前被阅读的文章都不足三篇，因此，文章 d3 对应的待定历史文章只有文章 d2 和文章 d2；相应的，文章 d2 对应的待定历史文章为文章 d1，而文章 d1 之前不存在其他被阅读的文章，所以针对文章 d1 不需要分析其对应的待定历史文章。

相应的，针对每篇历史文章，计算机设备需要分别计算该历史文章与该历史文章对应的至少一篇待定历史文章之间的注意力分数，从而该历史文章对应的注意力序列。参见图 6 中从上到下第一个箭头指向的部分图像，在图 6 中该部分，通过连线“-”连接两篇文章的标识来表示出两篇文章之间的注意力分数。如图 6 中所示，对于文章 d6，需要计算文章 d6 与文章 d5 之间的注意力分数，在图 6 中其表示为 d6-d5，注意力分数为 0.6；相应的，文章 d6 与文章 d4 之间的注意力分数 d6-d4 为 0.1；文章 d6 与文章 d3 之间的注意力分数为 d6-d3 为 0.3。对于其他文章 d5、d4、d3、d2 也类似，具体可以参见图 6 中“注意力分数序列”部分。

结合图 6 中从上到下的第二个箭头指向的部分图像中示出的了待定历史文章与历史文章之间的对应关系，由此可以看出，每篇文章可能会作为多篇其他文章的待定历史文章。在该部分中示出了作为待定历史文章的每篇文章所对应的多篇历史文章，及该待定历史文章与各篇历史文章之间的注意力分数。如，文章 d5 仅仅是文章 d6 对应的待定历史文章，而该文章 d5 与文章 d6 的注意力分数 d6-d5 可以从文章 d6 对应的注意力分数序列中查出，即文章 d5 与文章 d6 的注意力分数 d6-d5 为 0.6。相应的，文章 d5 对应的平均注意力分数为 0.6。

又如，文章 d4 作为文章 d6 以及文章 d5 的待定历史文章，其中，文章 d4 与文章 d6 的注意力分数 d6-d4 为 0.1；而文章 d4 与文章 d5 的注意力分数 d5-d4 为 0.4，由此可知，该待定历史文章 d4 对应的平均注意力分数为 0.25。

与此类似，由图 6 可以看出，得到作为待定历史文章的文章 d3 对应的平均注意力分数为 0.4；作为待定历史文章的文章 d2 对应的平均注意力 0.33；作为待定历史文章的文章 d1 对应的平均注意力分数为 0.56。

可见，对于所有的待定历史文章，文章 d4 对应的平均注意力分数最低，则将文章 d4 确定为推荐参考文章，以便后续以文章 d4 为依据，确定需要推荐给目标用户的候选推荐文章。

在另一种实现方式，计算机设备可以默认每篇待定历史文章对应着预设数量篇历史文章，如果某篇待定历史文章对应的历史文章的数量不足该预设数量，而预设数量与根据待定历史文章对应的历史文章的数量之间的差值，补充该差值个默认注意力程度（如，注意力分数）。然后，分别计算待定历史文章与预设数量篇历史文章之间的注意力程度的加和，并将加和后的注意力程度作为反映该待定历史文章具备触发推荐文章的可能性的依据。

举例说明，仍以图 6 为例，假设需要确定出每篇待定历史文章与 3 篇历史文章之间的注意力分数，并假设默认注意力分数为 0.5。在该图 6 中文章 d5 仅仅是文章 d6 对应的待定历史文章，因此，文章 d5 对应的历史文章的数量不足预设数量，即 3 篇。在该种情况下，文章 d5 对应的注意力程度的加和可以为文章 d5 与文章 d6 的注意力分数与两个默认注意力分数加和，具体为： $0.6+0.5+0.5$ 等于 1.6。而文章 b3 对应的历史文章有三篇，即为文章 b6、文章 b5 和文章 b4，在该种情况下，可以直接将文章 b3 分别与这两篇历史文章的注意力分数加和即可，即文章 b3 对应的加和后的注意力分数为： $0.3+0.5+0.4=1.2$ 。由此可以看出，由文章 b3 触发推荐文章的可能性，小于由文章 b5 触发推荐文章的可能性。

S305，计算机设备分别根据每篇该推荐参考文章的内容特征，从可推荐文章集合中，确定待推荐给该目标用户的至少一篇候选推荐文章。

其中，可推荐文章集合中包括能够推荐给该目标用户的多篇文章，如针对一个文章发布平台，该可推荐文章集合可以为文章发布平台所能发布的所有文章的集合；又如，该可推荐文章集合也可以是文章发布平台中所有用户阅读过的文章所构成的集合，当然，还可能有其他可能情况，在此不加以限制。

其中，针对每篇推荐参考文章，基于该推荐参考文章，计算机设备确

定待推荐给该目标用户的候选推荐文章的具体方式可以有多种。如，在一种可能的实现方式中，计算机设备可以依据推荐参考文章的内容特征，并结合文章协同推荐算法、基于内容的推荐算法或者基于序列的推荐算法等，从该可推荐文章集合中，确定出至少一篇候选推荐文章。

可选的，为了使得确定出的候选推荐文章能够与该待推荐参考文章的类型或者内容的相关程度更高，计算机设备还可以基于注意力的推荐策略，确定候选推荐文章。其中，基于注意力的推荐策略为：针对每篇推荐参考文章，可以依据该推荐参考文章的内容特征以及可推荐文章集合中每篇文章的内容特征，分别确定该推荐参考文章与该可推荐文章集合中每篇文章的注意力程度。相应的，可以从该可推荐文章集合中，选取出与该推荐参考文章的注意力程度较高的至少一篇候选推荐文章。

在确定候选推荐文章时，也可以确定出多篇候选推荐文章。例如，在计算可推荐文章集中每篇文章与至少一篇的推荐参考文章的第二注意力程度后，可以对得到的第二注意力程度进行排序，第二注意力程度越高则表示用户阅读推荐参考文章后，阅读该篇文章的可能性越大。若是从可推荐文章集合中确定出两篇候选推荐文章，则可以从可推荐文章集合中选择第二注意力程度最大的文章与第二注意力程度第二大的文章。

其中，计算机设备确定推荐参考文章与可推荐文章集合中各篇文章的注意力程度的方式可以参考前面确定注意力程度的相关介绍，如，可以将可推荐文章集合的各篇文章的文章向量，以及该推荐参考文章的文章向量输入到预先训练出的注意力模型中，以输出该推荐参考文章与该可推荐文章集合中各篇文章之间的注意力程度（如，注意力分数）所构成的注意力分数序列。当然，前面描述的通过其他方式确定注意力程度的方式也同样适用于此处，在此不再赘述。

可见，在本申请实施例中，对于用户阅读过的历史文章，计算机设备会分析用户在阅读该历史文章之前阅读的其他历史文章与该历史文章的注意力程度，由于两篇文章之间的注意力程度可以反映出用户阅读一篇文章的情况下，推荐用户阅读另一篇文章的可能性，因此，依据确定出的各

篇文章与其他文章之间的注意力程度，可以从阅读历史集合中确定出表征向目标用户推荐文章的可能性较低的历史文章，而该历史文章的类型及内容便不属于用户经常阅读到的文章类型及文章内容，因此，以该历史文章作为推荐参考文章，向用户推荐候选推荐文章，有利于提高确定出的候选推荐文章的多样性，进而有利于提高向用户推荐文章的多样性。

可以理解的是，基于推荐参考文章，确定出候选推荐文章之后，计算机设备可以直接将候选推荐文章作为需要推荐给该目标用户的文章。然而，确定出的考虑到候选推荐文章的数量有可能会相对较大；而且，在实际应用中，可能会配置多种推荐策略，并分别依据不同的推荐策略，确定与该推荐参考文章匹配的多篇候选推荐文章，也会使得确定出的候选推荐文章的数量较大，如果将这些大量的候选推荐文章都推荐给用户，很难实现精准向用户推荐用户感兴趣的文章，从而影响到推荐效果。

可选的，在依据可推荐文章集合中各文章与推荐参考文章之间的注意力程度，基于内容推荐算法，协同推荐算法等方式中的一种或者多种，确定出至少一个候选推荐文章之后，计算机设备还可以针对每篇候选推荐文章，分别计算历史阅读集合中每篇历史文章与该候选推荐文章的注意力程度，得到候选推荐文章与该历史阅读集合中多篇历史文章之间的注意力程度序列。然后，分别计算每篇候选推荐文章对应的注意力程度序列的信息熵，得到每篇候选推荐文章各自对应的信息熵。相应的，可以从该至少一篇候选推荐文章中，选取出信息熵较小的至少一篇候选推荐文章作为待推荐给该目标用户的至少一篇目标推荐文章。

可选地，在确定出每一篇候选推荐文章的信息熵后，可以对信息熵进行排序，信息熵越小，则对应的候选推荐文章越优先推送。如需要推送两篇目标推荐文章时，则对候选推荐文章的信息熵进行排序，并获取信息熵最小的候选推荐文章与信息熵第二小的候选推荐文章，将该两篇候选推荐文章作为目标推荐文章推送给用户。

可以理解的是，历史阅读集合中每篇历史文章与该候选推荐文章的注

意力程度可以反映出,由各篇历史文章触发阅读该候选推荐文章的可能性;而,该候选推荐文章与该历史阅读集合中各篇历史文章之间的注意力程度序列对应的信息熵可以衡量该历史阅读集合中各篇历史文章触发阅读到该候选推荐文章的稳定性或者说可靠性。可见,通过选取对应的信息熵较小的候选推荐文章作为待推荐的目标推荐文章,有利于在增加向用户推荐文章多样性的前提下,最大程度的保证向用户推荐用户感兴趣的文章,以实现基于用户进行个性化推荐。

计算机设备确定出至少一篇目标推荐文章之后,还可以确定这至少一篇目标推荐文章的排序,以便后续文章发布平台按照该排序输出该至少一篇目标推荐文章。

如,参见图7,其示出了本申请实施例一种文章推荐方法的又一种流程示意图,本实施例的方法从计算机设备侧描述,该方法可以包括:

S701,计算机设备获取待分析的目标用户的历史阅读集合。

其中,该历史阅读集合包括:目标用户在不同阅读时刻阅读过的多篇历史文章。

S702,针对每篇历史文章,依据该历史文章的内容特征,计算机设备确定该历史文章的文章向量,并依据该历史阅读集合中阅读时刻在该历史文章的阅读时刻之前的至少一篇待定历史文章各自的内容特征,分别确定每篇待定历史文章的文章向量。

S703,针对每篇历史文章,计算机设备将该历史文章对应的至少一篇待定历史文章的文章向量以及该历史文章的内容向量输入到预先训练出的注意力模型中,以通过该注意力模型输出注意力分数序列。

其中,该注意力分数序列包括该至少一篇待定历史文章分别与该历史文章之间的注意力分数。待定历史文章与该历史文章的注意力分数可以反映在用户阅读待定历史文章的情况下,推荐用户阅读该历史文章的可能性。

S704,针对每篇该待定历史文章,计算机设备根据该待定历史文章分别到至少一篇历史文章的注意力分数,确定该待定历史文章到至少一篇历史文章的平均注意力分数。

S705, 计算机设备从多篇该历史文章对应的多篇待定历史文章中, 选出平均注意力分数较低的至少一篇待定历史文章作为推荐参考文章。

以上步骤 S701 到 S705 可以参见前面的相关介绍, 在此不再赘述。其中, 步骤 S704 和 S705 仅仅是基于待定历史文章与相应的历史文章之间的注意力分数, 确定推荐参考文章的一种实现方式, 对于前面实施例提到的其他方式, 也同样适用于本实施例, 在此不再赘述。

可以理解的是, 以上步骤 S701 到 S705 实际上是基于注意力模型的选取策略, 即, 基于注意力模型, 从目标用户的阅读历史集合中选取推荐参考文章。但是可以理解的是, 在实际应用中, 基于注意力模型, 来选取推荐参考文章可以仅仅是选取待推荐文章的一种选取策略, 因此, 在计算机设备中除了配置基于注意力模型的选取策略之后, 还可以同时配置其他一种或多种选取策略, 如基于用户画像的选取策略, 即, 基于目标用户的用户画像, 从该阅读历史集合中选出一篇或多篇历史文章作为推荐参考文章。相应的, 计算机设备可以分别通过配置的多种选取策略, 分别选取出的推荐参考文章, 并执行后续的操作。

如, 参见图 8, 本申请中文章推荐方法的一种实现原理示意图。由图 8 可以看出, 用于选取推荐参考文章的推荐策略可以有多种。相应的, 依据目标用户阅读的阅读历史集合以及用户的目标用户的用户画像, 并分别采用多种不同的推荐策略, 从阅读历史集合中确定推荐参考文章, 这样, 基于每种推荐策略可以均可以选出一篇推荐参考文章。

S706, 针对每篇推荐参考文章, 计算机设备依据该推荐参考文章的内容特征以及可推荐文章集合中每篇文章的内容特征, 分别确定可推荐文章集合中每篇文章与推荐参考文章的第二注意力程度。

其中, 可推荐文章集合中任意一篇文章与推荐参考文章之间第二注意力程度表征用户在阅读推荐参考文章的情况下, 推荐用户阅读该可推荐文章集合中该文章的可能性。

S707, 从该推荐文章集合中, 计算机设备选取与该推荐参考文章的第二注意力程度较高的至少一篇候选推荐文章。

可以理解的是，文章与推荐参考文章的注意力程度越高，则说明文章与该推荐参考文章的类型及内容的关联性越高，则该文章属于用户阅读过的但是又不经常阅读过的文章类型，将该文章选取为候选推荐文章，则既能贴合用户的阅读需求，又有可能增加用户阅读的文章的多样性。

在以上步骤 S706 至 S707 中实际上是基于注意力的推荐策略，来确定候选推荐文章为例进行说明，但是可以理解的通过其他的推荐策略，来确定候选推荐文章的方式也同样适用于本实施例。

可以理解的是，在实际应用中，计算机设备基于推荐参考文章，确定候选推荐文章时，可以仅仅设置一种推荐策略，如上面提到的基于注意力的推荐策略。在实际应用中，可能会设置多种推荐策略，如，设置基于注意力的推荐策略的同时，还会设置基于内容的推荐策略，以及基于协同的推荐策略等等。如图 8，推荐策略也可以配置有多种。相应的，计算机设备可以分别以各篇推荐参考文章为基准，并分别多种推荐策略，从可推荐文章集合中确定出相应的至少一篇候选推荐文章；或者是，综合多种推荐策略，来确定出至少一篇候选推荐文章。如图 8 所示，依据多种推荐策略以及确定出的推荐参考文章，可以从推荐文章集合中选取候选推荐文章，以便执行后续的初选操作。

S708，针对每篇候选推荐文章，计算机设备分别计算该历史阅读集合中每篇历史文章与该候选推荐文章的第三注意力程度，得到候选推荐文章与该历史阅读集合中多篇历史文章之间的注意力程度序列。

其中，该注意力程度序列中包括候选推荐文章分别与该历史阅读集合中多篇历史文章的第三注意力程度。

S709，计算机设备分别计算每篇候选推荐文章对应的注意力程度序列的信息熵，得到每篇候选推荐文章各自对应的信息熵。

其中，该注意力程度序列对应的信息熵可以反映出该候选推荐文章与历史阅读集合中各篇历史文章之间注意力程度的分布情况。候选推荐文章对应的注意力程度序列的信息熵越小，则说明根据用户阅读的历史阅读集合，确定出用户对该候选推荐文章的感兴趣程度越高。

其中，计算注意力程度序列的信息熵的方式可以有多种，本申请对此不加以限制。

为了便于理解，以一种计算信息熵的方式进行介绍。如，假设候选推荐文章与历史文章集合的注意力程度序列为一个注意力分数序列，其表示为 $A_i = \{a_1, a_2, \dots, a_i\}$ ，则该注意力分数序列对应的信息熵 $H(A)$ 可以通过如下公式计算得到：

$$H(A) = - \sum_{i=1}^I a_i \log a_i \quad (\text{公式四});$$

其中，公式四中的 a_i 为注意力分数序列 A_i 中的注意力分数。

S710，从该至少一篇候选推荐文章中，选取出信息熵较小的至少一篇候选推荐文章作为待推荐给该目标用户的至少一篇目标推荐文章。

计算机设备通过选取对应的信息熵较小的候选推荐文章作为目标推荐文章，可以使得筛选出的目标推荐文章在满足多样性的前提下，更符合用户的兴趣。

可以理解的是，以上步骤 S708 到 S710 是基于注意力模型，从候选推荐文章中，确定最终筛选推荐的目标推荐文章的方式，即，基于注意力模型的初选策略。但是可以理解的是，通过其他初选策略，从多篇候选推荐文章中最终筛选出至少一篇目标推荐文章也同样适用于本实施例。如，初选策略可以为基于用户画像，来从多篇候选推荐文章中选取与用户画像最为匹配的至少一篇目标推荐文章。

在实际应用中，还可以在基于注意力模型的初选策略的基础上，综合其他初选策略，以最终从候选推荐文章中筛选出至少一篇目标推荐文章。如图 8 所示，计算机设备中可以配置多种初选策略。相应的，计算机设备可以分别基于该多种初选策略，从候选推荐文章中确定出至少一篇目标推荐文章；或者是，结合多种初选策略，来综合确定出至少一篇候选推荐文章。

S711，针对任意一篇推荐参考文章，计算机设备确定该至少一篇目标推荐文章分别与推荐参考文章的第四注意力程度。

其中，第四注意力程度反映用户在阅读该推荐参考文章的情况下，推荐用户阅读该目标推荐文章的可能性。

S712，依据该至少一篇目标推荐文章分别与推荐参考文章之间的第四注意力程度，计算机设备确定该至少一篇目标推荐文章的推荐顺序。

如，目标推荐文章对应的第四注意力程度越高，该目标推荐文章的排序越靠前。特别的，当存在多篇推荐参考文章的情况下，可以依据目标推荐文章分别与各篇推荐参考文章的第四注意力程度，来确定该目标推荐文章的第四注意力程度的综合值，并根据各个目标推荐文章的该综合值，来确定推荐顺序。

其中，目标推荐文章对应的第四注意力程度越高，则在用户阅读该推荐参考文章的情况下，选择阅读该目标推荐文章的可能性越大，因此，对应的第四注意力程度较高的目标推荐文章的排在较为靠前的顺序上，越有利于用户点击该目标推荐文章。

可以理解的是，计算机设备确定出该至少一篇目标推荐文章的推荐顺序之后，可以将该至少一篇目标推荐文章的标识以及推荐顺序发送给文章发布服务器。相应的，在目标用户访问文章发布服务器，如，登录文章发布服务器阅读文章等，则该文章发布服务器可以依据该推荐顺序，向该目标用户展现该至少一篇目标推荐文章，以使得目标用户可以看到更为广泛类型且较为符合自身兴趣的文章。

可以理解的是，第二注意力程度，第三注意力程度、第四注意力程度仅仅是为了便于区分不同文章之间的注意力程度，但是其求取过程都可以参见前面介绍的用于确定文章之间的注意力程度的方式，如，利用预先训练得到的注意力模型来确定以上的第二注意力程度、第三注意力程度及第四注意力程度，具体可以参见前面实施例的相关介绍，在此不再赘述。

另一方面，本申请还提供了一种文章推荐方法，可选地，如图9所示，上述文章推荐方法包括：

S902，计算机设备获取待分析的目标用户的历史阅读集合，历史阅读

集合包括：目标用户在不同阅读时刻阅读过的多篇历史文章；

S904，计算机设备针对多篇历史文章中的每篇历史文章，确定阅读时刻在每篇历史文章之前的每篇待定历史文章与每篇历史文章的第一注意力程度，第一注意力程度反映在用户阅读待定历史文章的情况下，推荐用户阅读每篇历史文章的可能性；

S906，计算机设备依据每篇待定历史文章分别与至少一篇历史文章的第一注意力程度，从多篇待定历史文章中选取出至少一篇推荐参考文章；

S908，计算机设备分别根据每篇推荐参考文章，从可推荐文章集合中，确定待推荐给目标用户的至少一篇候选推荐文章。

以下将上述文章推荐方法应用到具体的推荐过程中进行说明。计算机设备为手机。如图 10 所示，图 10 为一种可选的用户的浏览界面，用户可以使用终端 1002 登录客户端浏览文章，图 10 中显示有客户端根据现有的推荐方法已经推荐的文章 1004。用户可以通过点击相关的文章标题或者链接阅读一些文章。在用户的阅读过程中，会产生一些阅读记录，阅读记录包括阅读的文章与阅读时刻。将阅读记录作为用户的历史阅读集合。历史阅读集合中包括了用户阅读过的多篇历史文章。

在获取到上述多篇历史文章之后，手机会获取多篇历史文章中每一篇历史文章的待定历史文章。待定历史文章为阅读时刻在上述每一篇历史文章之前的文章。例如，以多篇历史文章包括“我们是冠军”与“赛程表”、

“主演表”、“古装”为例，四篇历史文章的阅读顺序从先到后为：“我们是冠军”、“赛程表”、“主演表”、“古装”。“古装”的待定历史文章为位于“古装”之前的三篇历史文章；“主演表”的待定历史文章为位于“主演表”之前的两篇历史文章；“赛程表”的待定历史文章为位于“赛程表”之前的一篇历史文章；“我们是冠军”没有待定历史文章。手机确定出每一篇历史文章的待定历史文章之后，确定每一篇待定历史文章与对应的历史文章的第一注意力程度。如表（1），表 1 中的横线“-”表示不包含第一注意力程度。

表（1）

第一注意力程度	我们是冠军	赛程表	主演表	古装
我们是冠军	-	0.9	0.4	0.2
赛程表	-	-	0.1	0.3
主演表	-	-	-	0.8
古装	-	-	-	-

确定第一注意力程度的方法可以采用上述提到的各种方法，在此不再赘述。可见，在用户阅读“我们是冠军”后，阅读“赛程表”的可能性较高，而阅读“主演表”的可能性较低。而对于每一篇待定历史文章，如“我们是冠军”，可以计算“我们是冠军”和与其对应的至少一篇的历史文章的第一注意力程度。如上述计算的“我们是冠军”与“赛程表”的第一注意力程度为 0.9、“我们是冠军”与“演员表”的第一注意力程度为 0.4、“我们是冠军”与“古装”的第一注意力程度为 0.2，得到一个平均值为 0.5。采用同样的方法，得到每一篇待定历史文章的第一注意力程度的平均值，“赛程表”的第一注意力平均值为 0.2，“主演表”的第一注意力程度为 0.8。

手机将多篇待定历史文章中，第一注意力平均值最低的“赛程表”确定为推荐参考文章，然后根据推荐参考文章“赛程表”从可推荐文章集合中推荐出一篇候选推荐文章。如，可推荐文章集合中包括“赛程时间”、“导演表”，则可以将“赛程时间”作为候选推荐文章，在用户获取推荐文章时，将候选推荐文章推荐给用户。如图 11 所示，图 11 中用户通过下拉的方法请求推荐文章，如图 12 所示，手机将“赛程时间”推荐给用户，显示在推荐区 1202 中。

另一方面，本申请还提供了一种文章推荐装置。如，参见图 13，其示出了本申请一种文章推荐装置一个实施例的组成结构示意图，本实施例的装置可以应用于前面的计算机设备，该装置可以包括：

历史获取单元 1301，设置为获取待分析的目标用户的历史阅读集合，历史阅读集合包括：目标用户在不同阅读时刻阅读过的多篇历史文章；

第一注意分析单元 1302, 设置为针对每篇历史文章, 依据历史文章的内容特征以及历史阅读集合中阅读时刻在历史文章的阅读时刻之前的至少一篇待定历史文章的内容特征, 分别确定每篇待定历史文章与历史文章的第一注意力程度, 第一注意力程度反映在用户阅读待定历史文章的情况下, 推荐用户阅读历史文章的可能性;

推荐参考确定单元 1303, 设置为依据得到的多篇待定历史文章分别与至少一篇历史文章的第一注意力程度, 从多篇待定历史文章中, 选取出触发向目标用户推荐文章的可能性相对较低的至少一篇待定历史文章作为推荐参考文章;

候选推荐确定单元 1304, 设置为分别根据每篇推荐参考文章的内容特征, 从可推荐文章集合中, 确定待推荐给目标用户的至少一篇候选推荐文章。

可选的, 第一注意分析单元可以包括:

第一注意分析子单元, 设置为依据历史文章的内容特征以及历史阅读集合中阅读时刻在历史文章的阅读时刻之前的至少一篇待定历史文章的内容特征, 并利用预先训练得到的注意力模型, 分别确定每篇待定历史文章与历史文章的注意力分数;

其中, 注意力模型为利用多个用户的历史文章样本集合以及每个历史文章样本集合中每篇历史文章样本各自对应的已标注的注意力分数序列训练的到的, 其中, 历史文章样本的注意力分数序列包括: 用户阅读该历史文章样本之前阅读的多篇历史文章样本分别与该历史文章样本的注意力分数。

可选的, 第一注意力分析子单元可以包括:

第一向量确定子单元, 设置为依据历史文章的内容特征, 确定历史文章的文章向量;

第二向量确定子单元, 设置为依据历史阅读集合中阅读时刻在历史文章的阅读时刻之前的至少一篇待定历史文章各自的内容特征, 分别确定每篇待定历史文章的文章向量;

注意分数分析子单元, 设置为根据历史文章的文章向量以及至少一篇待定历史文章各自的文章向量, 并利用注意力模型, 分别确定每篇待定历史文章与历史文章的注意力分数。

在一种可能的实现方式中, 推荐参考确定单元, 包括:

平均注意分析子单元, 设置为针对每篇待定历史文章, 根据待定历史文章分别与至少一篇历史文章的第一注意力程度, 确定待定历史文章到至少一篇历史文章的平均注意力程度;

参考筛选子单元, 设置为从多篇历史文章对应的多篇待定历史文章中, 选取出平均注意力程度较低的至少一篇待定历史文章作为推荐参考文章。

在一种可能的实现方式中, 候选推荐确定单元, 包括:

第二注意分析单元, 设置为针对每篇推荐参考文章, 依据推荐参考文章的内容特征以及可推荐文章集合中每篇文章的内容特征, 分别确定可推荐文章集合中每篇文章与推荐参考文章的第二注意力程度, 第二注意力程度反映用户在阅读推荐参考文章的情况下, 推荐用户阅读可推荐文章集合中文章的可能性;

候选推荐选取单元, 设置为从可推荐文章集合中, 选取出与推荐参考文章的第二注意力程度较高的至少一篇候选推荐文章。

如, 参见图 14, 其示出了本申请的文章推荐装置的又一种组成结构示意图, 本实施例的文章推荐装置与前面图 14 所示装置的不同之处在于:

该装置除了包括, 历史获取单元 1401, 第一注意分析单元 1402, 推荐参考确定单元 1403 和候选推荐确定单元 1404 之外, 还包括:

第三注意分析单元 1405, 设置为在候选推荐确定单元确定出至少一篇候选推荐文章之后, 针对每篇候选推荐文章, 分别计算历史阅读集合中每篇历史文章与候选推荐文章的第三注意力程度, 得到候选推荐文章与历史阅读集合中多篇历史文章之间的注意力程度序列, 第三注意力程度反映用户在阅读历史阅读集合中的历史文章的情况下, 推荐用户阅读候选推荐文章的可能性;

信息熵计算单元 1406, 设置为分别计算每篇候选推荐文章对应的注意力程度序列的信息熵, 得到每篇候选推荐文章各自对应的信息熵;

目标推荐确定单元 1407, 设置为从至少一篇候选推荐文章中, 选取出信息熵较小的至少一篇候选推荐文章作为待推荐给目标用户的至少一篇目标推荐文章。

可选的, 该装置还可以包括:

第四注意分析单元 1408, 设置为分别确定推荐参考文章与至少一篇目标推荐文章的第四注意力程度, 第四注意力程度反映用户在阅读推荐参考文章的情况下, 推荐用户阅读目标推荐文章的可能性;

排序确定单元 1409, 设置为依据推荐参考文章分别与至少一篇目标推荐文章的第四注意力程度, 确定至少一篇目标推荐文章的推荐顺序。

其中, 历史获取单元 1401, 第一注意分析单元 1402, 推荐参考确定单元 1403 和候选推荐确定单元 1404 具体可以参见前面实施例的相关介绍, 在此不在赘述。

根据本申请的实施例的又一方面, 还提供了一种存储介质, 该存储介质中存储有计算机程序, 其中, 该计算机程序被设置为运行时执行上述任一项方法实施例中的步骤。

可选地, 在本实施例中, 上述存储介质可以被设置为存储用于执行以下步骤的计算机程序:

S1, 获取待分析的目标用户的历史阅读集合, 历史阅读集合包括: 目标用户在不同阅读时刻阅读过的多篇历史文章;

S2, 针对每篇历史文章, 依据历史文章的内容特征以及历史阅读集合中阅读时刻在历史文章的阅读时刻之前的至少一篇待定历史文章的内容特征, 分别确定每篇待定历史文章与历史文章的第一注意力程度, 第一注意力程度反映在用户阅读待定历史文章的情况下, 推荐用户阅读历史文章的可能性;

S3, 依据得到的多篇待定历史文章分别与至少一篇历史文章的第一注意力程度, 从多篇待定历史文章中, 选取出触发向目标用户推荐文章的可能性相对较低的至少一篇待定历史文章作为推荐参考文章;

S4, 分别根据每篇推荐参考文章的内容特征, 从可推荐文章集合中, 确定待推荐给目标用户的至少一篇候选推荐文章。

或者, 在本实施例中, 上述存储介质可以被设置为存储用于执行以下步骤的计算机程序:

S1, 计算机设备获取待分析的目标用户的历史阅读集合, 历史阅读集合包括: 目标用户在不同阅读时刻阅读过的多篇历史文章;

S2, 计算机设备针对多篇历史文章中的每篇历史文章, 确定阅读时刻在每篇历史文章之前的每篇待定历史文章与每篇历史文章的第一注意力程度, 第一注意力程度反映在用户阅读待定历史文章的情况下, 推荐用户阅读每篇历史文章的可能性;

S3, 计算机设备依据每篇待定历史文章分别与至少一篇历史文章的第一注意力程度, 从多篇待定历史文章中选取出至少一篇推荐参考文章;

S4, 计算机设备分别根据每篇推荐参考文章, 从可推荐文章集合中, 确定待推荐给目标用户的至少一篇候选推荐文章。

可选地, 在本实施例中, 本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通程序来指令终端设备相关的硬件来完成, 该程序可以存储于一计算机可读存储介质中, 存储介质可以包括: 闪存盘、只读存储器(Read-Only Memory, ROM)、随机存取器(Random Access Memory, RAM)、磁盘或光盘等。

需要说明的是, 本说明书中的各个实施例均采用递进的方式描述, 每个实施例重点说明的都是与其他实施例的不同之处, 各个实施例之间相同相似的部分互相参见即可。对于装置类实施例而言, 由于其与方法实施例

基本相似，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

最后，还需要说明的是，在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括要素的过程、方法、物品或者设备中还存在另外的相同要素。

对所公开的实施例的上述说明，使本领域技术人员能够实现或使用本申请。对这些实施例的多种修改对本领域技术人员来说将是显而易见的，本文中所定义的一般原理可以在不脱离本申请的精神或范围的情况下，在其它实施例中实现。因此，本申请将不会被限制于本文所示的这些实施例，而是要符合与本文所公开的原理和新颖特点相一致的最宽的范围。

以上仅是本申请的可选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本申请原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也应视为本申请的保护范围。

工业实用性

在本申请实施例中，对于用户阅读过的历史文章，会分析用户在阅读该历史文章之前阅读的其他历史文章与该历史文章的注意力程度，由于两篇文章之间的注意力程度可以反映出用户阅读一篇文章的情况下，推荐用户阅读另一篇文章的可能性，因此，依据确定出的各篇文章与其他文章之间的注意力程度，可以从阅读历史集合中确定出表征向目标用户推荐文章的可能性较低的历史文章，而该历史文章的类型及内容便不属于用户经常阅读到的文章类型及文章内容，因此，以该历史文章作为推荐参考文章，

向用户推荐候选推荐文章，有利于提高确定出的候选推荐文章的多样性，进而有利于提高向用户推荐文章的多样性。

权 利 要 求 书

1. 一种文章推荐方法，包括：

计算机设备获取待分析的目标用户的历史阅读集合，所述历史阅读集合包括：所述目标用户在不同阅读时刻阅读过的多篇历史文章；

所述计算机设备针对所述多篇历史文章中的每篇历史文章，确定阅读时刻在所述每篇历史文章之前的每篇待定历史文章与所述每篇历史文章的第一注意力程度，所述第一注意力程度反映在用户阅读所述待定历史文章的情况下，推荐用户阅读所述每篇历史文章的可能性；

所述计算机设备依据每篇所述待定历史文章分别与至少一篇历史文章的第一注意力程度，从多篇所述待定历史文章中选取出至少一篇推荐参考文章；

所述计算机设备分别根据每篇所述推荐参考文章，从可推荐文章集合中，确定待推荐给所述目标用户的至少一篇候选推荐文章。

2. 根据权利要求 1 所述的的文章推荐方法，其中，所述计算机设备针对所述多篇历史文章中的每篇历史文章，确定阅读时刻在所述每篇历史文章之前的每篇待定历史文章与所述每篇历史文章的第一注意力程度包括：

所述计算机设备依据所述每篇历史文章的内容特征以及所述历史阅读集合中阅读时刻在所述每篇历史文章的阅读时刻之前的至少一篇待定历史文章的内容特征，并利用预先训练得到的注意力模型，分别确定每篇所述待定历史文章与所述每篇历史文章的注意力分数；

其中，所述注意力模型为利用多个用户的历史文章样本集合以及每个所述历史文章样本集合中每篇历史文章样本各自对应的已标注的注意力分数序列训练得到的，其中，历史文章样本的注意力分数序列包括：用户阅读该历史文章样本之前阅读的多篇历史文章样本分别与该历史文章样本的注意力分数。

3. 根据权利要求 2 所述的的文章推荐方法，其中，所述计算机设备利用预先训练得到的注意力模型，分别确定每篇所述待定历史文章与所述历史

文章的注意力分数，包括：

所述计算机设备依据所述历史文章的内容特征，确定所述历史文章的文章向量；

所述计算机设备依据所述历史阅读集合中阅读时刻在所述历史文章的阅读时刻之前的至少一篇待定历史文章各自的内容特征，分别确定每篇待定历史文章的文章向量；

所述计算机设备根据所述历史文章的文章向量以及所述至少一篇待定历史文章各自的文章向量，并利用所述注意力模型，分别确定每篇待定历史文章与所述历史文章的注意力分数。

4. 根据权利要求 1 至 3 任一项所述的文章推荐方法，其中，所述计算机设备依据每篇所述待定历史文章分别与至少一篇历史文章的第一注意力程度，从多篇所述待定历史文章中选取出至少一篇推荐参考文章包括：

所述计算机设备针对每篇所述待定历史文章，根据所述待定历史文章分别与至少一篇历史文章的第一注意力程度，确定所述待定历史文章到所述至少一篇历史文章的平均注意力程度；

所述计算机设备对所述待定历史文章的所述平均注意力程度由小到大进行排序，并获取与排序结果中前 X 个平均注意力程度对应的 X 篇待定历史文章，并将所述 X 篇待定历史文章确定为推荐参考文章，其中，所述 X 根据所述多篇历史文章的数量确定。

5. 根据权利要求 1 所述的文章推荐方法，其中，所述计算机设备分别根据每篇所述推荐参考文章，从可推荐文章集合中，确定待推荐给所述目标用户的至少一篇候选推荐文章，包括：

所述计算机设备针对每篇推荐参考文章，依据所述推荐参考文章的内容特征以及可推荐文章集合中每篇文章的内容特征，分别确定所述可推荐文章集合中每篇文章与所述推荐参考文章的第二注意力程度，所述第二注意力程度反映用户在阅读所述推荐参考文章的情况下，推荐用户阅读所述可推荐文章集合中文章的可能性；

所述计算机设备对所述可推荐文章集合中的每篇文章的第二注意

力程度由大到小进行排序，得到排序结果，并从所述可推荐文章集合中获取与所述排序结果中前 Y 个第二注意力程度对应的 Y 篇文章，将所述 Y 篇文章确定为候选推荐文章，其中，所述 Y 根据所述可推荐文章集合中的文章的数量确定。

6. 根据权利要求 1 或 5 所述的文章推荐方法，其中，所述计算机设备确定出至少一篇候选推荐文章之后，还包括：

所述计算机设备针对每篇候选推荐文章，分别计算所述历史阅读集合中每篇历史文章与所述候选推荐文章的第三注意力程度，得到所述候选推荐文章与所述历史阅读集合中多篇历史文章之间的注意力程度序列，所述第三注意力程度反映用户在阅读所述历史阅读集合中的历史文章的情况下，推荐用户阅读所述候选推荐文章的可能性；

所述计算机设备分别计算每篇候选推荐文章对应的注意力程度序列的信息熵，得到每篇候选推荐文章各自对应的信息熵；

所述计算机设备对所述每篇候选推荐文章的信息熵由小到大进行排序，并获取排序结果中前 Z 个信息熵所对应的 Z 篇候选推荐文章，将所述 Z 篇候选推荐文章确定为待推荐给所述目标用户的目标推荐文章，其中，所述 Z 根据所述候选推荐文章的数量确定。

7. 根据权利要求 6 所述的文章推荐方法，其中，在所述计算机设备对所述每篇候选推荐文章的信息熵由小到大进行排序，并获取排序结果中前 Z 个信息熵所对应的 Z 篇候选推荐文章，将所述 Z 篇候选推荐文章确定为待推荐给所述目标用户的目标推荐文章之后，还包括：

所述计算机设备分别确定所述推荐参考文章与至少一篇目标推荐文章的第四注意力程度，所述第四注意力程度反映用户在阅读所述推荐参考文章的情况下，推荐用户阅读所述目标推荐文章的可能性；

依据所述推荐参考文章分别与所述至少一篇目标推荐文章的第四注意力程度，确定所述至少一篇目标推荐文章的推荐顺序。

8. 一种文章推荐装置，应用在计算机设备中，包括：

历史获取单元，设置为获取待分析的目标用户的历史阅读集合，

所述历史阅读集合包括：所述目标用户在不同阅读时刻阅读过的多篇历史文章；

第一注意分析单元，设置为针对所述多篇历史文章中的每篇历史文章，确定阅读时刻在所述每篇历史文章之前的每篇待定历史文章与所述每篇历史文章的第一注意力程度，所述第一注意力程度反映在用户阅读所述待定历史文章的情况下，推荐用户阅读所述每篇历史文章的可能性；

推荐参考确定单元，设置为依据每篇所述待定历史文章分别与至少一篇历史文章的第一注意力程度，从多篇所述待定历史文章中选取至少一篇推荐参考文章；

候选推荐确定单元，设置为分别根据每篇所述推荐参考文章，从可推荐文章集合中，确定待推荐给所述目标用户的至少一篇候选推荐文章。

9. 根据权利要求 8 所述的文章推荐装置，其中，所述第一注意分析单元，包括：

第一注意分析子单元，设置为依据所述每篇历史文章的内容特征以及所述历史阅读集合中阅读时刻在所述每篇历史文章的阅读时刻之前的至少一篇待定历史文章的内容特征，并利用预先训练得到的注意力模型，分别确定每篇所述待定历史文章与所述每篇历史文章的注意力分数；

其中，所述注意力模型为利用多个用户的历史文章样本集合以及每个所述历史文章样本集合中每篇历史文章样本各自对应的已标注的注意力分数序列训练的得到的，其中，历史文章样本的注意力分数序列包括：用户阅读该历史文章样本之前阅读的多篇历史文章样本分别与该历史文章样本的注意力分数。

10. 根据权利要求 8 或 9 所述的文章推荐装置，其中，所述推荐参考确定单元，包括：

平均注意分析子单元，设置为针对每篇所述待定历史文章，根据

所述待定历史文章分别与至少一篇历史文章的第一注意力程度，确定所述待定历史文章到所述至少一篇历史文章的平均注意力程度；

参考筛选子单元，设置为对所述待定历史文章的所述平均注意力程度由小到大进行排序，并获取与排序结果中前 X 个平均注意力程度对应的 X 篇待定历史文章，并将所述 X 篇待定历史文章确定为推荐参考文章，其中，所述 X 根据所述多篇历史文章的数量确定。

11. 根据权利要求 8 所述的装置，其中，所述候选推荐确定单元，包括：

第二注意分析单元，设置为针对每篇推荐参考文章，依据所述推荐参考文章的内容特征以及可推荐文章集合中每篇文章的内容特征，分别确定所述可推荐文章集合中每篇文章与所述推荐参考文章的第二注意力程度，所述第二注意力程度反映用户在阅读所述推荐参考文章的情况下，推荐用户阅读所述可推荐文章集合中文章的可能性；

候选推荐选取单元，设置为对所述可推荐文章集合中每一篇文章的第二注意力程度由大到小进行排序，得到排序结果，并从所述可推荐文章集合中获取与所述排序结果中前 Y 个第二注意力程度对应的 Y 篇文章，将所述 Y 篇文章确定为候选推荐文章，其中，所述 Y 根据所述可推荐文章集合中的文章的数量确定。

12. 根据权利要求 8 或 11 所述的装置，其中，还包括：

第三注意分析单元，设置为在所述候选推荐确定单元确定出至少一篇候选推荐文章之后，针对每篇候选推荐文章，分别计算所述历史阅读集合中每篇历史文章与所述候选推荐文章的第三注意力程度，得到所述候选推荐文章与所述历史阅读集合中多篇历史文章之间的注意力程度序列，所述第三注意力程度反映用户在阅读所述历史阅读集合中的历史文章的情况下，推荐用户阅读所述候选推荐文章的可能性；

信息熵计算单元，设置为分别计算每篇候选推荐文章对应的注意力程度序列的信息熵，得到每篇候选推荐文章各自对应的信息熵；

目标推荐确定单元，设置为对所述每篇候选推荐文章的信息熵由

小到大进行排序，并获取排序结果中前 Z 个信息熵所对应的 Z 篇候选推荐文章，将所述 Z 篇候选推荐文章确定为待推荐给所述目标用户的目标推荐文章，其中，所述 Z 根据所述候选推荐文章的数量确定。

13. 根据权利要求 12 所述的推荐装置，其中，还包括：

第四注意分析单元，设置为分别确定所述推荐参考文章与至少一篇目标推荐文章的第四注意力程度，所述第四注意力程度反映用户在阅读所述推荐参考文章的情况下，推荐用户阅读所述目标推荐文章的可能性；

排序确定单元，设置为依据所述推荐参考文章分别与所述至少一篇目标推荐文章的第四注意力程度，确定所述至少一篇目标推荐文章的推荐顺序。

14. 一种计算机设备，包括：

处理器和存储器；

其中，所述处理器用于执行所述存储器中存储的计算机程序；

所述存储器用于存储计算机程序，所述计算机程序至少用于：

获取待分析的目标用户的历史阅读集合，所述历史阅读集合包括：所述目标用户在不同阅读时刻阅读过的多篇历史文章；

针对所述多篇历史文章中的每篇历史文章，确定阅读时刻在所述每篇历史文章之前的每篇待读历史文章与所述每篇历史文章的第一注意力程度，所述第一注意力程度反映在用户阅读所述待读历史文章的情况下，推荐用户阅读所述每篇历史文章的可能性；

依据每篇所述待读历史文章分别与至少一篇历史文章的第一注意力程度，从多篇所述待读历史文章中选取出至少一篇推荐参考文章；

分别根据每篇所述推荐参考文章，从可推荐文章集合中，确定待推荐给所述目标用户的至少一篇候选推荐文章。

15. 一种存储介质，所述存储介质中存储有计算机程序，所述计算机程序被运行时，执行所述权利要求 1 至 8 任一项中所述的推荐方法。

说明书附图

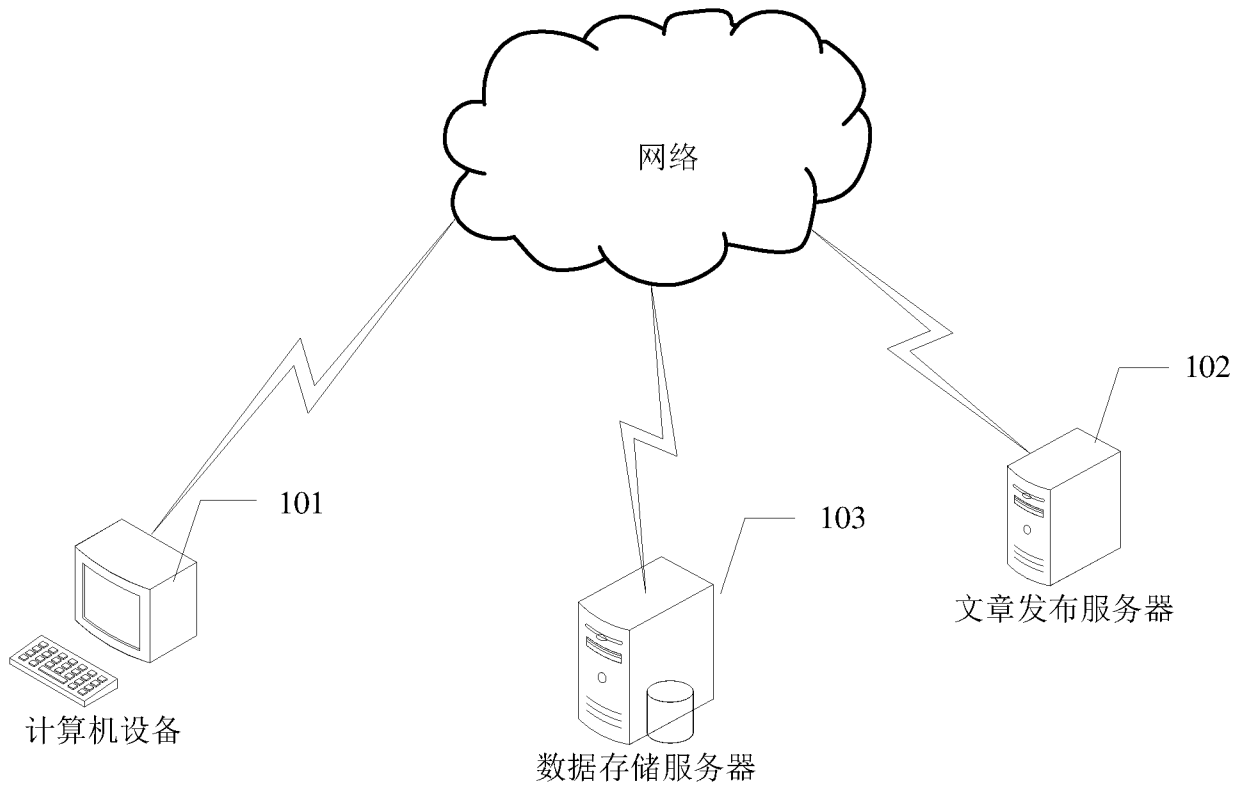


图 1

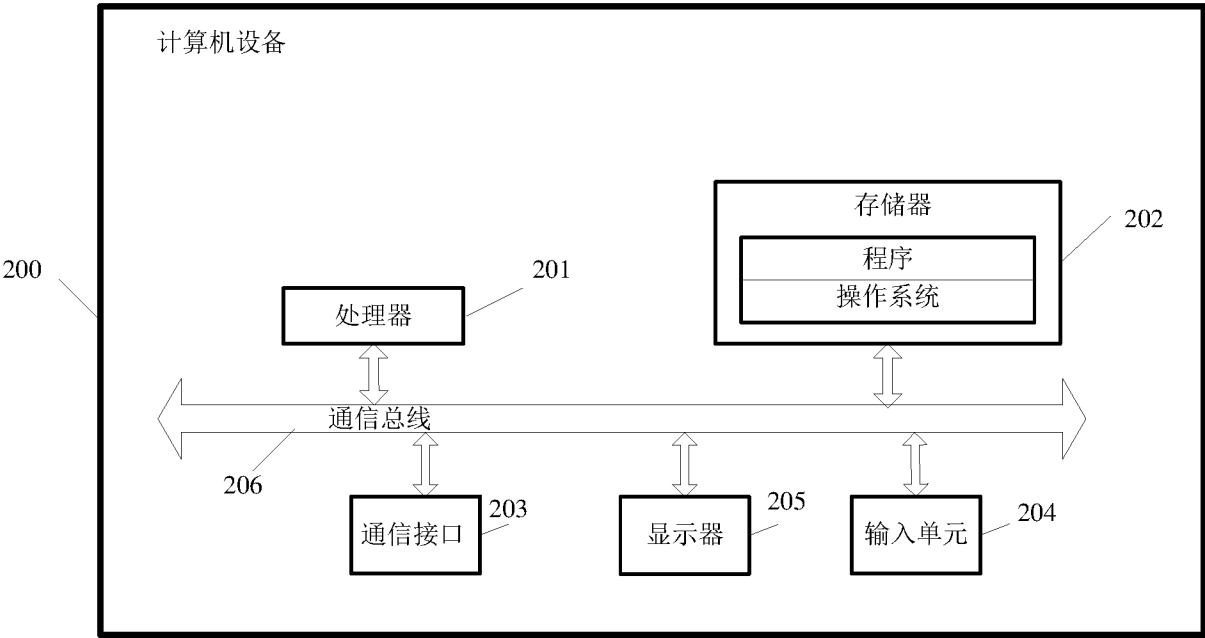


图 2

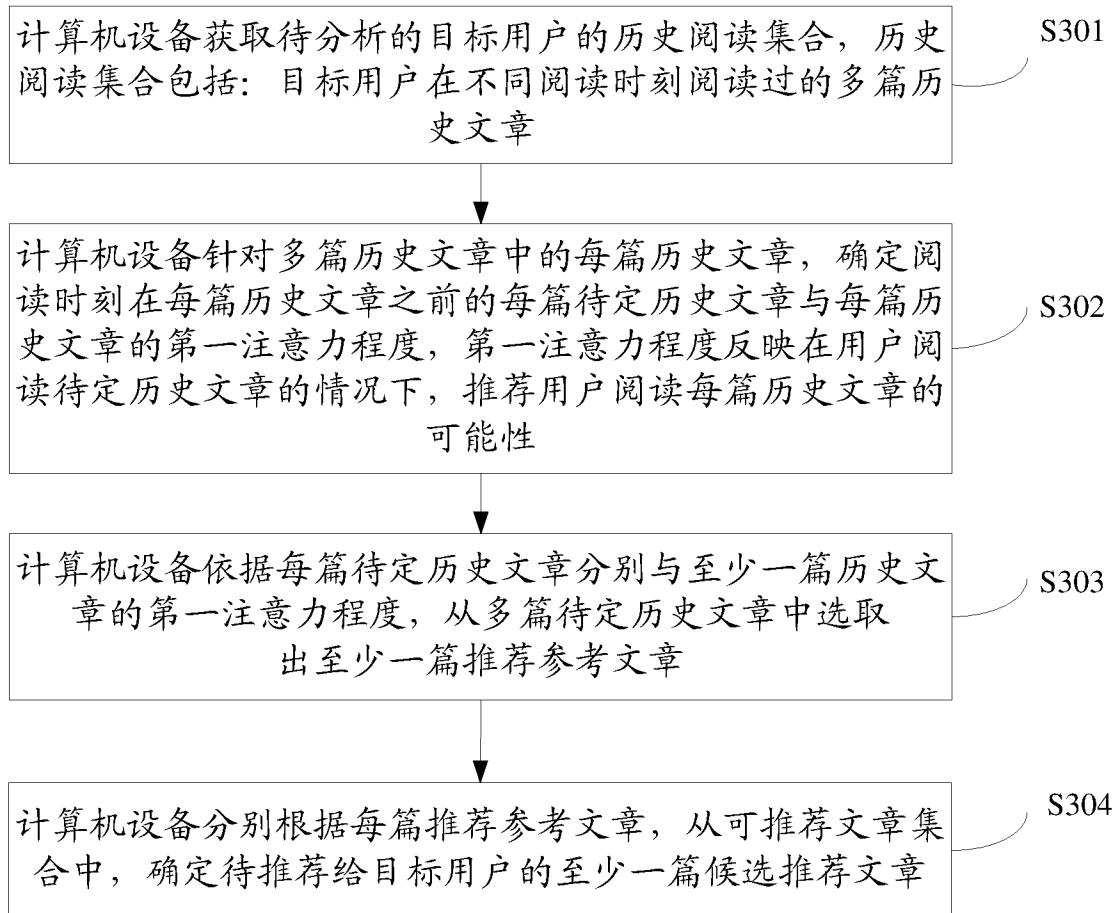


图 3

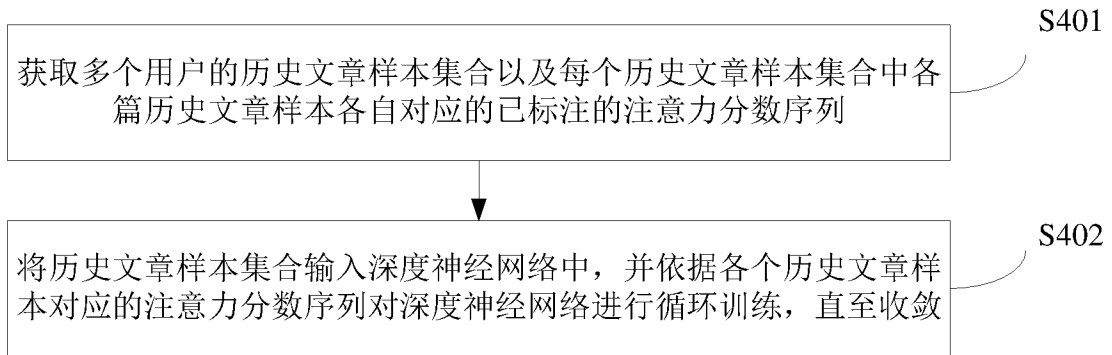


图 4

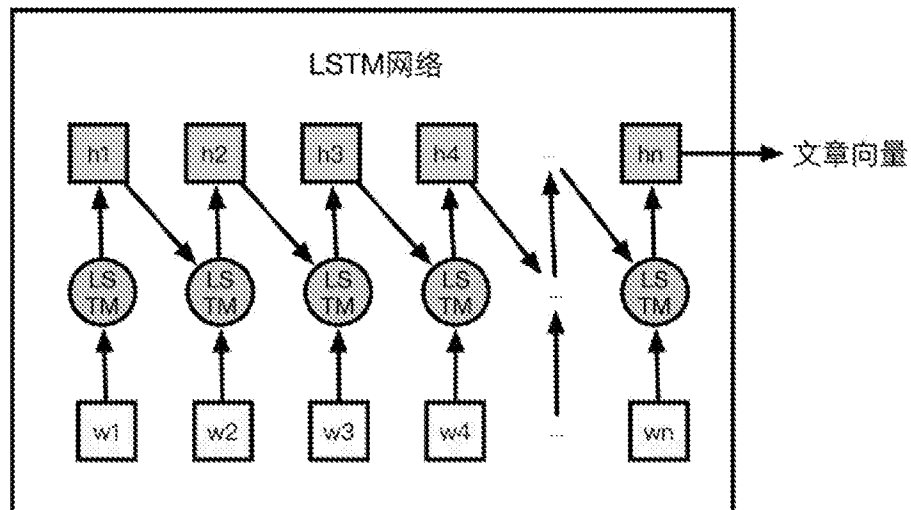


图 5

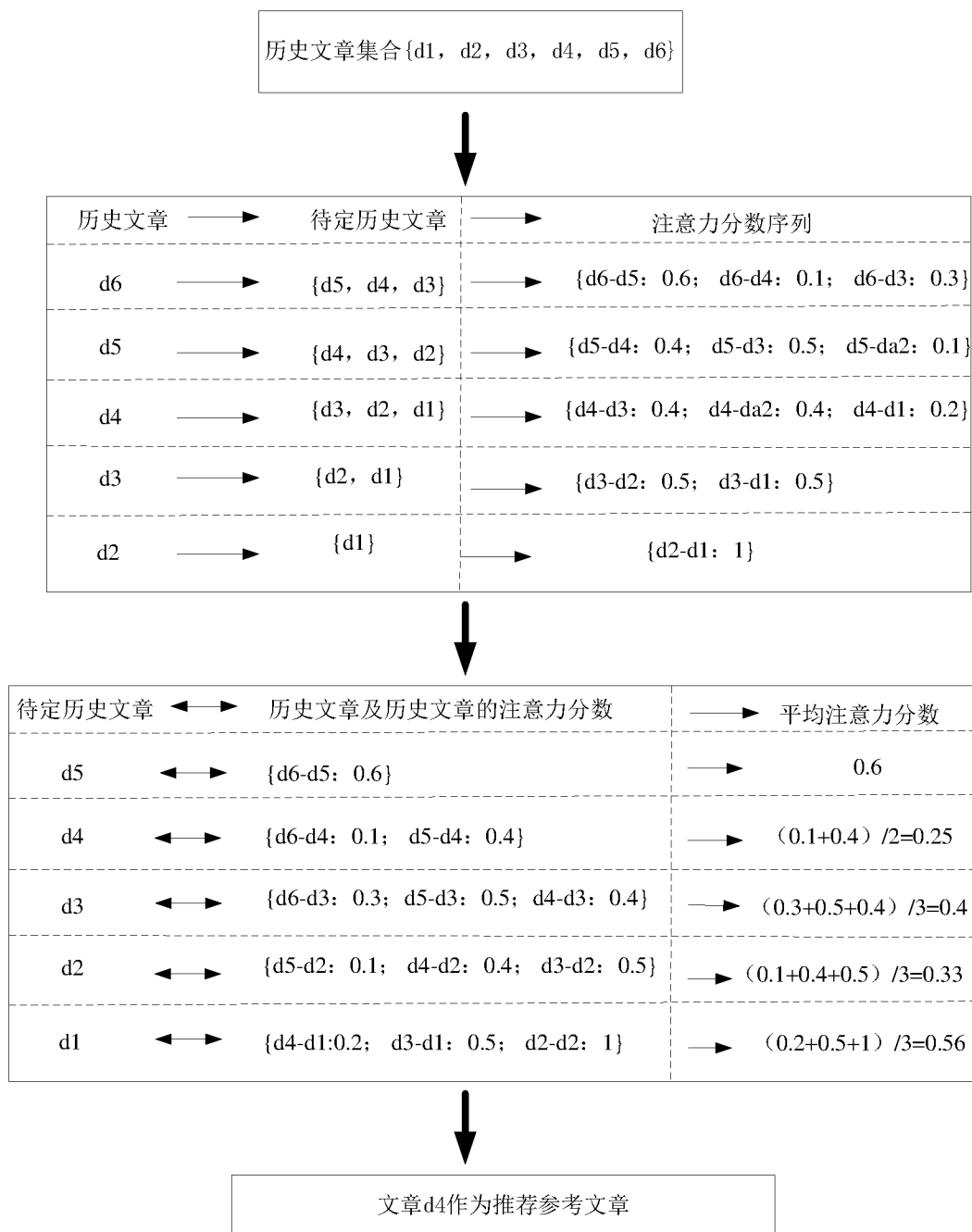


图 6

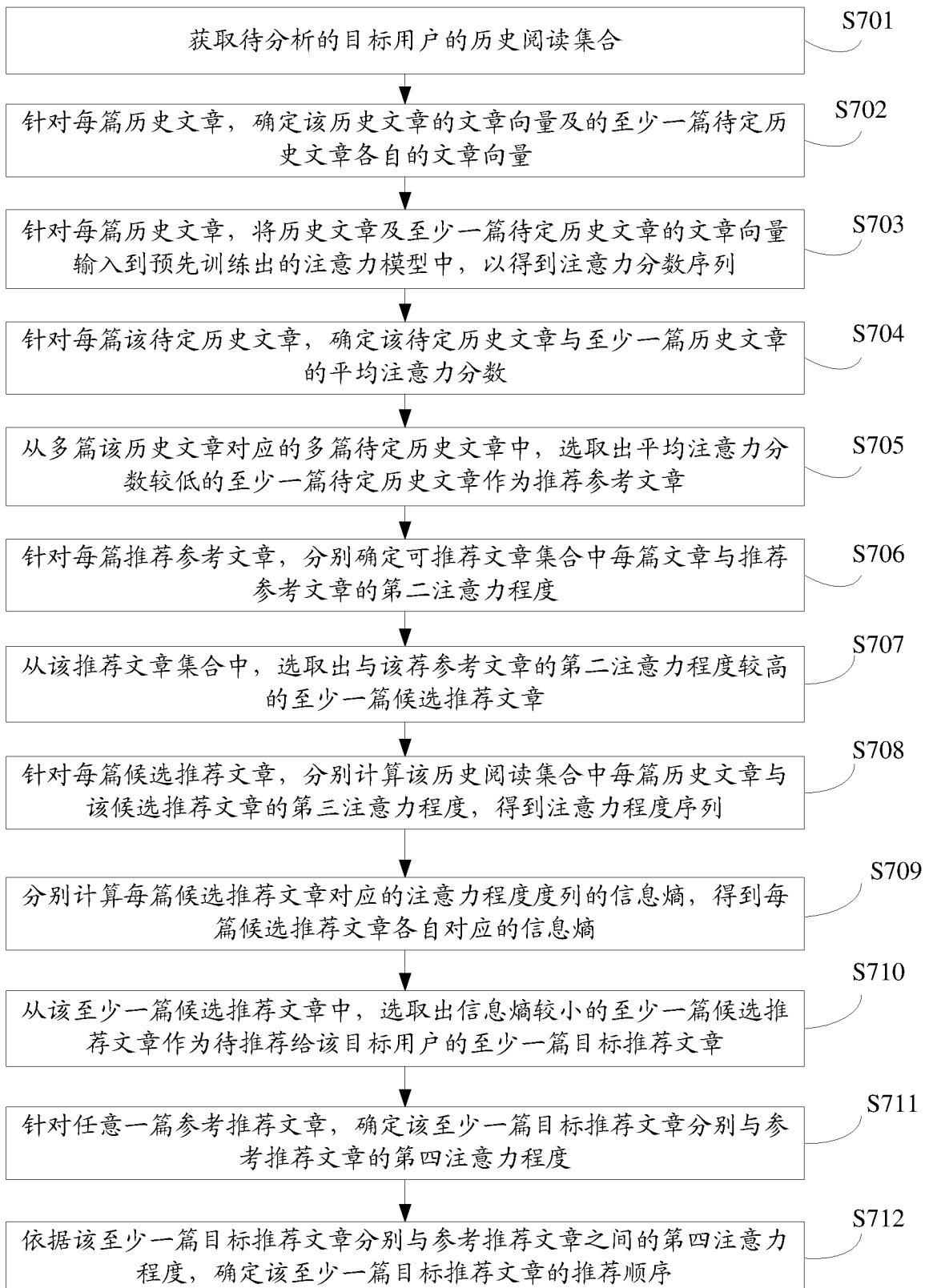


图 7

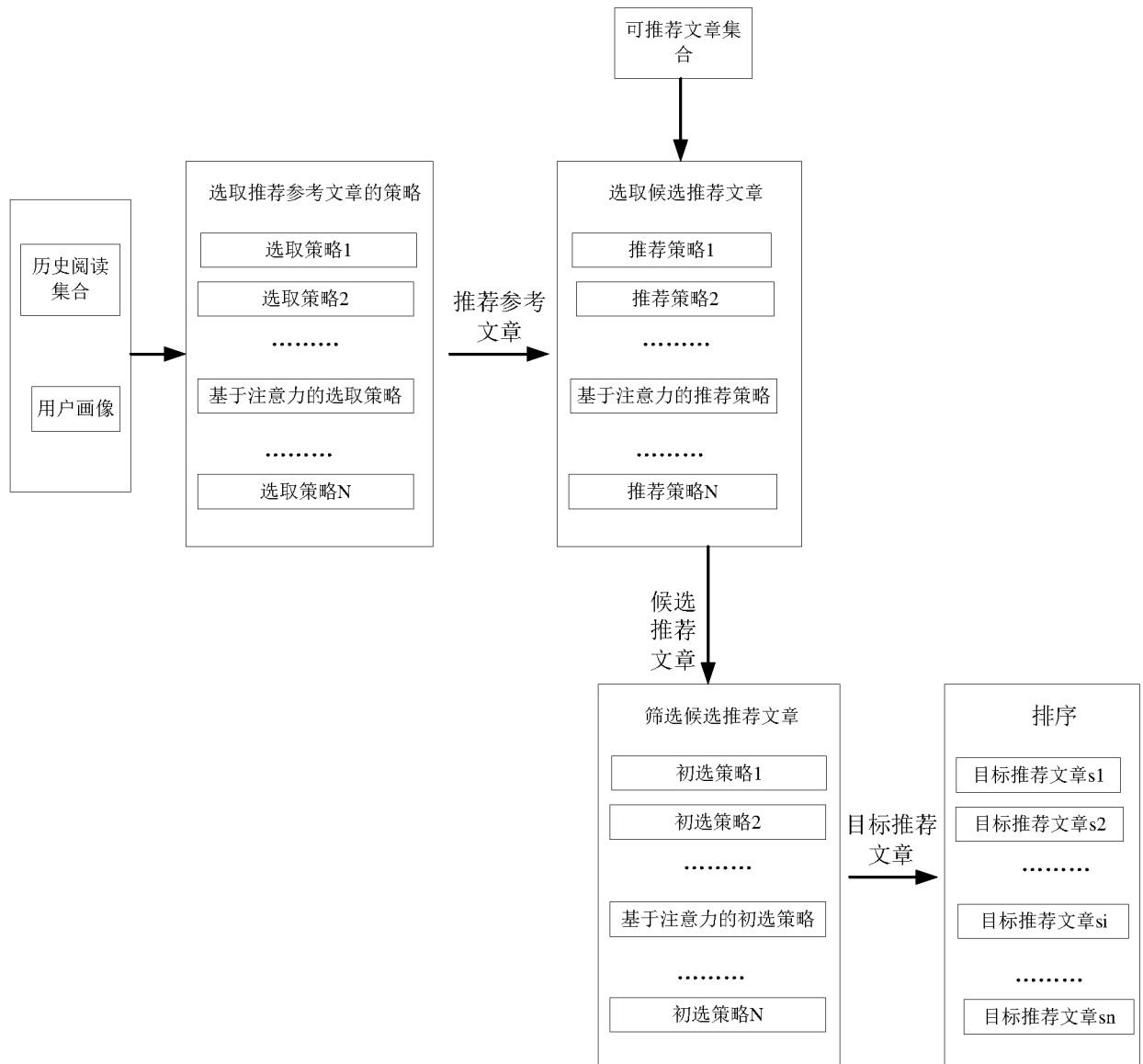


图 8

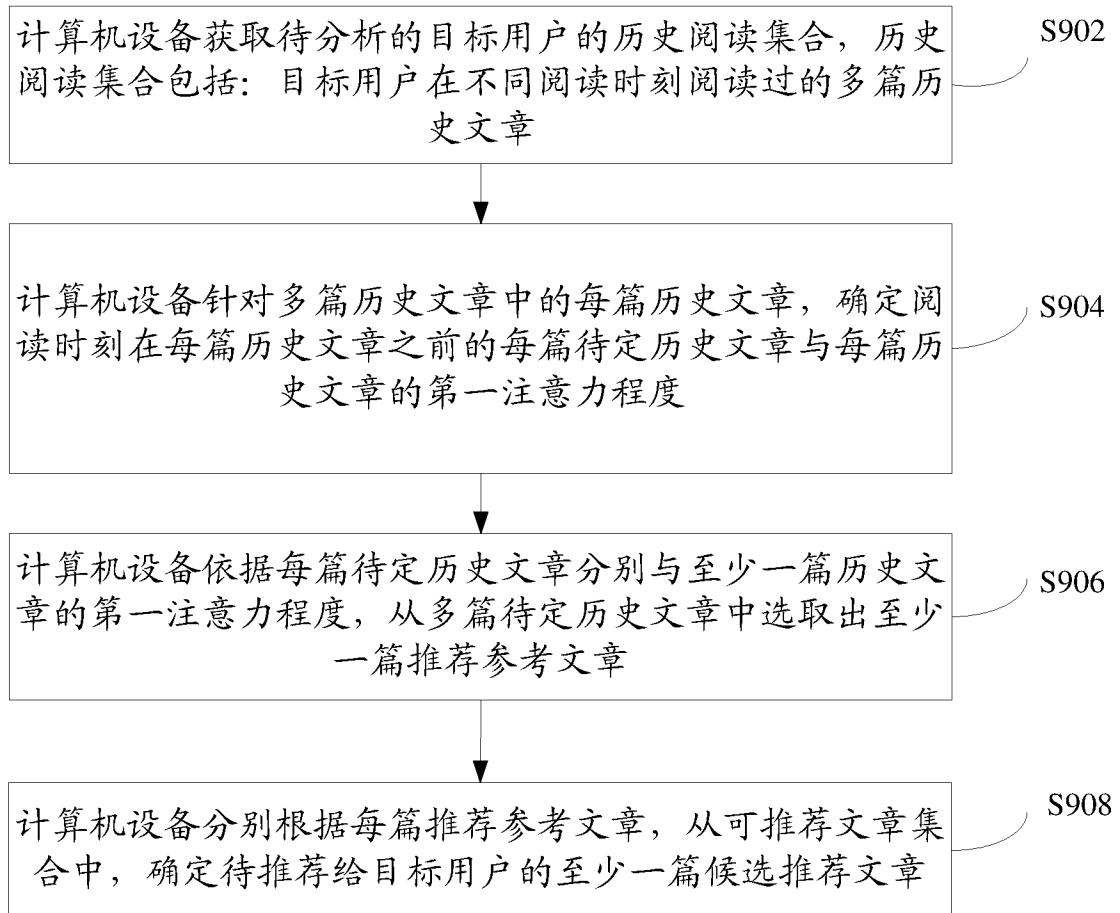


图 9

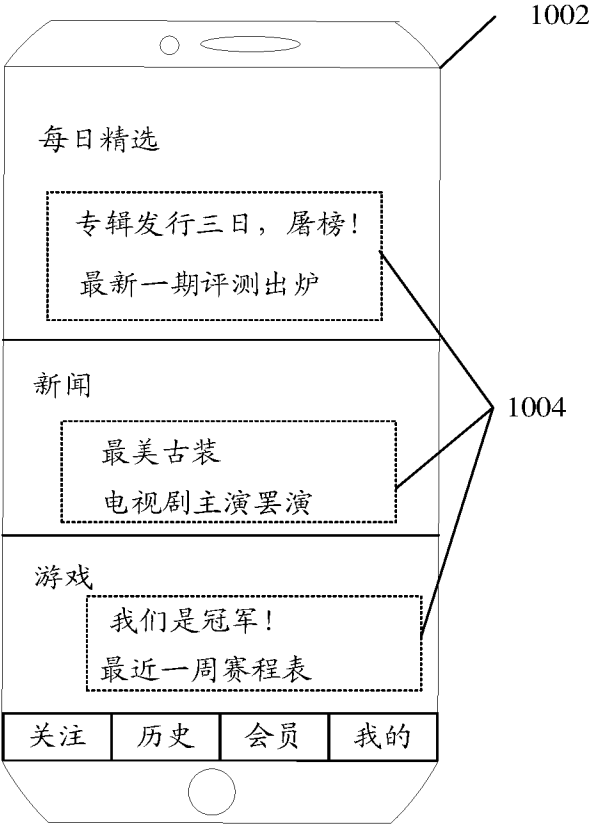


图 10

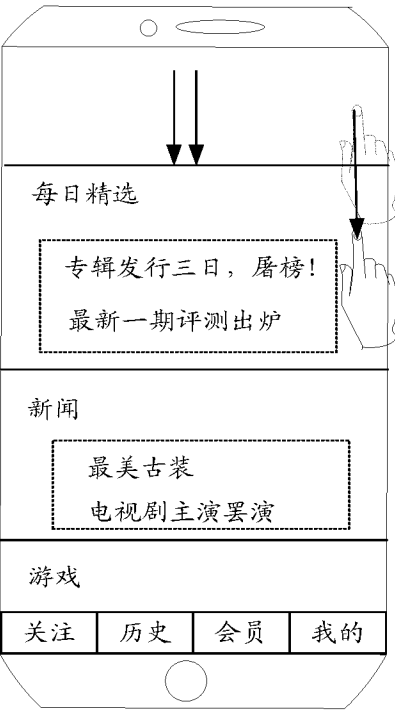


图 11

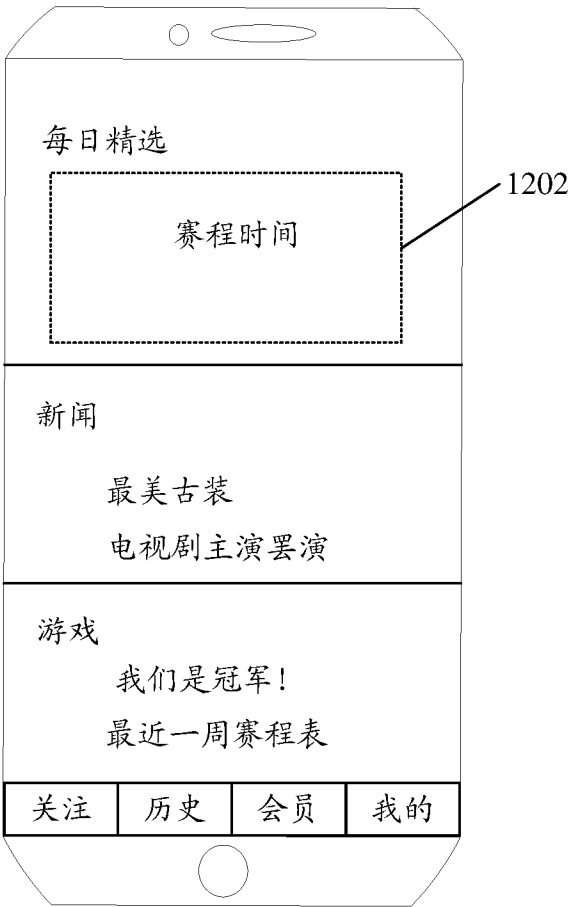


图 12

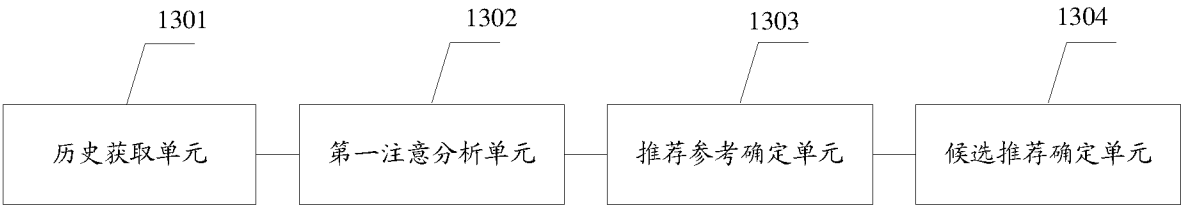


图 13

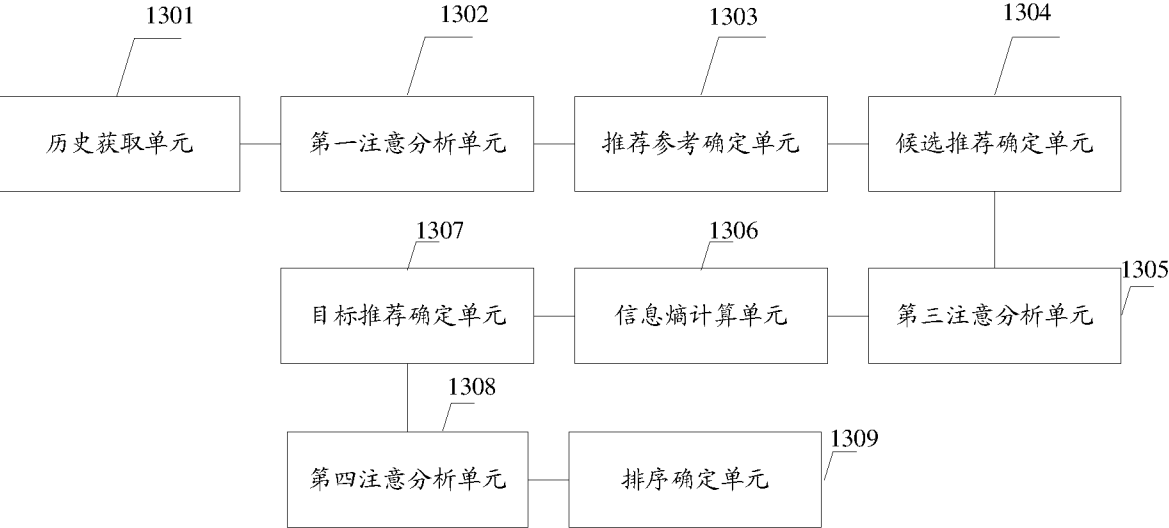


图 14

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/086374

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/00(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNTXT; CNABS; CNKI; TWTXT; TWABS; USTXT; VEN; EPTXT; WOTXT: 推荐, 推送, 历史, 阅读, 可能性, 多样, 模型, 训练, recommend, push, history, read, possibility, diversity, diversify, model, train

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 103559262 A (BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 05 February 2014 (2014-02-05) description, paragraphs 21-39	1-15
A	CN 103678620 A (STATE GRID CORPORATION OF CHINA ET AL.) 26 March 2014 (2014-03-26) description, paragraphs 5-67	1-15
A	CN 103488714 A (HANGZHOU EASTCOM-BUPT INFORMATION TECHNOLOGY LTD.) 01 January 2014 (2014-01-01) entire document	1-15
A	US 2018107645 A1 (SKYWRITERRX INC.) 19 April 2018 (2018-04-19) entire document	1-15
A	US 2015120712 A1 (YAHOO INC.) 30 April 2015 (2015-04-30) entire document	1-15

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

22 July 2019

Date of mailing of the international search report

12 August 2019

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/086374

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	103559262	A	05 February 2014	CN	103559262	B	05 October 2016
CN	103678620	A	26 March 2014	CN	103678620	B	15 February 2017
CN	103488714	A	01 January 2014	CN	103488714	B	18 January 2017
US	2018107645	A1	19 April 2018	None			
US	2015120712	A1	30 April 2015	US	2017308537	A1	26 October 2017
				US	9703783	B2	11 July 2017

国际检索报告

国际申请号

PCT/CN2019/086374

A. 主题的分类

G06F 17/00 (2019.01) i

按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献 (标明分类系统和分类号)

G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))

CNTXT;CNABS;CNKI;TWTXT;TWABS;USTXT;VEN;EPTXT;WOTXT:推荐, 推送, 历史, 阅读, 可能性, 多样, 模型, 训练, recommend, push, history, read, possibility, diversity, diversify, model, train

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 103559262 A (北京邮电大学) 2014年 2月 5日 (2014 - 02 - 05) 说明书第21-39段	1-15
A	CN 103678620 A (国家电网公司等) 2014年 3月 26日 (2014 - 03 - 26) 说明书第5-67段	1-15
A	CN 103488714 A (杭州东信北邮信息技术有限公司) 2014年 1月 1日 (2014 - 01 - 01) 全文	1-15
A	US 2018107645 A1 (SKYWRITER INC) 2018年 4月 19日 (2018 - 04 - 19) 全文	1-15
A	US 2015120712 A1 (YAHOO INC) 2015年 4月 30日 (2015 - 04 - 30) 全文	1-15

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2019年 7月 22日

国际检索报告邮寄日期

2019年 8月 12日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

杨牛

电话号码 86-(20)-28950388

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/086374

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	103559262	A	2014年 2月 5日	CN 103559262 B	2016年 10月 5日
CN	103678620	A	2014年 3月 26日	CN 103678620 B	2017年 2月 15日
CN	103488714	A	2014年 1月 1日	CN 103488714 B	2017年 1月 18日
US	2018107645	A1	2018年 4月 19日	无	
US	2015120712	A1	2015年 4月 30日	US 2017308537 A1	2017年 10月 26日
				US 9703783 B2	2017年 7月 11日